A Publication of the American Psychological Association

Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

ARTICLES

Maternal Permissiveness toward Aggression and Subsequent TAT Aggression ... 1
Donald Weatherley

Externalization in Perceptual Defense ... 6
David W. Lewit, A. Richard Breyer, and Alan H. Leiman

Dream Reports from Different Stages of Sleep ... 14
W. David Foulkes

The Competition of Affective Responses in Human Subjects ... 26
Gerald A. Mendelsohn

The Effects of Competition on Reaction Time and Palmar Skin Conductance ... 32
Russell M. Church

A Relationship between Threat in the Manifest Content of Dreams and Active-Passive Behavior in Psychotics ... 41
James L. Fromo, Jerre Osterweil, and Ivan Boszormenyi-Nagy

Motivation of Chronic Schizophrenics by Information about Their Abilities in a Group Situation ... 48
James C. Diggory and Arnold Loeb

Changes in Children's Verbal Behavior as a Function of Social Approval and Manifest Anxiety ... 53
Vinton Rowley and R. Dewayne Keller

CRITIQUE AND NOTES

Teaching Common Associations to Schizophrenics ... 58
Robert Sommer, Gwynneth Witney, and Humphry Osmond

Open- and Closed-Mindedness and the Ability to Differentiate Source and Message 61
Frederic A. Powell

Memorizing by Individuals and Groups: A Test of the Pooling-of-Ability Model 64
Ronald A. Hoppe

Convergent Validation of Risk Taking Measures ... 68
Paul Slovic

Social Schemas and the Reconstruction of Social Object Displays from Memory 71
James L. Kuethe

Volume 65 JULY 1962 Number 1

Volume 65 July 1962 Number 1

Journal of Abnormal and Social Psychology

Editor: DANIEL KATZ, *University of Michigan*

Associate Editor: ROBERT B. ZAJONC, *University of Michigan*

CONSULTING EDITORS

JOSEPH B. ADELSON
University of Michigan

GORDON W. ALLPORT
Harvard University

NORMAN H. ANDERSON
University of California, Los Angeles

LEONARD BERKOWITZ
University of Wisconsin

JACK W. BREHM
Duke University

URIE BRONFENBRENNER
Cornell University

DONALD R. BROWN
Bryn Mawr College

DONALD T. CAMPBELL
Northwestern University

LAUNOR F. CARTER
System Development Corporation

DWIGHT W. CHAPMAN, JR.
Vassar College

MORTON DEUTSCH
Bell Telephone Laboratories

SEYMOUR FESHBACH
University of Pennsylvania

DONALD W. FISKE
University of Chicago

NATHANIEL L. GAGE
University of Illinois

NICHOLAS HOBBS
Peabody College for Teachers

MURRAY HORWITZ
New York University

IRVING L. JANIS
Yale University

HAROLD H. KELLEY
University of California, Los Angeles

HERBERT C. KELMAN
Harvard University

DANIEL J. LEVINSON
Harvard Medical School

DAVID C. MCCLELLAND
Harvard University

ROBERT B. MACLEOD
Cornell University

HAROLD L. RAUSH
National Institute of Mental Health

NEVITT SANFORD
Stanford University

IRVING SARNOFF
Western Reserve University

WILLIAM A. SCOTT
University of Colorado

JANET TAYLOR SPENCE
Veterans Administration Hospital, Iowa City

JOSEPH VEROFF
University of Michigan

JUSTIN L. WEISS
Massachusetts Mental Health Center

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued monthly, two volumes per year. The subscription price per volume is $10.00, or $20.00 annually (foreign, $20.50). Single issues are $2.00.

Original manuscripts should be sent to the Editor:

DANIEL KATZ
University of Michigan
Box 1248
Ann Arbor, Michigan

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Acting Production Manager

JUDY SMITH
Editorial Assistant

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street, N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth Street N.W., Washington 6, D. C.

Second-class postage paid at Lancaster, Pa., and at additional mailing offices

Copyright © 1962 by the American Psychological Association, Inc.

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 1–5

MATERNAL PERMISSIVENESS TOWARD AGGRESSION AND SUBSEQUENT TAT AGGRESSION

DONALD WEATHERLEY

University of Colorado

A number of theorists working within the framework of a Hullian-derived learning theory (e.g., Bandura & Walters, 1959; Sears, Maccoby, & Levin, 1957), have argued that the extent to which a parent is permissive toward a child's aggression is an important determinant of that child's subsequent aggressiveness. According to this approach, aggression is a learned response. An individual's readiness to aggress is assumed to depend in part on the extent to which aggressive behavior has been successful in removing frustrations in the past. Such success constitutes one way in which aggression is reinforced. When a parent is relatively permissive toward a child's aggression it may be expected that the child will more frequently exhibit aggressive behavior and thereby increase the likelihood that his aggressive behavior will be associated with reinforcement.

But it is not only through its influence upon aggressive tendencies per se that degree of maternal permissiveness can affect subsequent aggressive behavior. When a mother permits very little aggression by restraining, punishing, or redirecting the child's aggressive behavior when it occurs, she must communicate to the child disapproval of such behavior, however subtle this communication might be. Consequently, internal and external cues relevant to aggression may come to evoke anxiety responses in the child which function to inhibit the direct expression of aggression. The greater the mother's disapproval, the stronger should be the internal restraints against aggression which develop in the child.

The effects of parental attitudes toward aggression are presumed to extend into adulthood. In the study reported here, two groups of college girls, who differ in the extent to which their mothers reported a permissive attitude toward childhood aggression, were compared with respect to their aggressive responses in TAT productions, under conditions of high and low aggression arousal and high and low cue relevance for aggression. An extension of the Miller approach-avoidance model was involved in the derivation of the predictions which were made. This paradigm has been used by a number of investigators (e.g., Clark, 1955; Scott, 1956; Weatherley, 1961) who have found that the experimental arousal of socially disapproved motives can lead to a diminution of expression of those motives in thematic test protocols. This finding suggests that the cues relevant to the arousal of a motive may also arouse an anxiety response which acts to depress the expression of the motive to a level below that obtained under neutral conditions.

In the present study a formulation based on the Miller goal-gradient model led to the expectation that experimental arousal of aggression would have a differential effect upon the TAT aggression of subjects differing in the degree of maternal permissiveness for aggression they experienced as children.

Figure 1 depicts the model used. The net aggressive response strength for either group of subjects at any point along the horizontal axis is the difference between the height of the appropriate approach gradient and the matching avoidance gradient at that point. As can be seen, the subjects treated relatively permissively as children—High Maternal Permissiveness (MP) subjects—may be expected to express more TAT aggression at Point A than the subjects whose mothers treated them relatively nonpermissively in childhood—Low MP subjects. This is true also at Point B and at all points in between. Furthermore, as arousal and/or cue relevance is increased from A to B, the High MP subjects should show an increase in TAT aggression, while the Low MP subjects should show a decrease.

The placement of points along the hypothetical dimension of arousal and/or cue

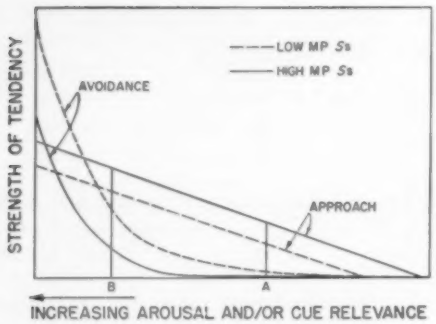

FIG. 1. Model depicting the aggressive approach and avoidance gradients of High and Low MP subjects.

relevance in Figure 1 is, of course, arbitrary; it is impossible to determine on an a priori basis what experimental operations and material will precisely correspond to Points A and B. In order to provide a broad sample of possible points along this dimension, two levels of aggression arousal and two levels of aggressive cue value of TAT cards were used.

Each subject told stories to TAT cards of high and low cue relevance for aggression under conditions of either high or low aggression arousal. Thus the variables of maternal permissiveness, degree of arousal, and cue value of TAT cards were ordered in a $2 \times 2 \times 2$ factorial design.

On the basis of the model depicted in Figure 1 two predictions were made: High MP subjects will express more TAT aggression in general than will Low MP subjects; High MP subjects will show a greater increase in TAT aggression as the cues and arousal conditions for aggression are increased than will Low MP subjects.

In addition to the TAT measure, subjects were asked to rate their subjective feelings during the experiment on several dimensions. No predictions were made concerning these measures; they were obtained to help evaluate the effectiveness of the aggression arousing situation and to supplement the information obtained from the TAT data.

METHOD

Procedure

The data reported in this paper were obtained from 100 female undergraduate students who volunteered to participate in the study. The data were collected in small group sessions. Half of the subjects served under High Arousal conditions: They were subjected to an aggression arousing situation immediately following which a TAT measure of aggression and ratings of the subjects' subjective feelings were obtained.

Aggression arousal was accomplished by making highly insulting and depreciating comments to the subjects in the course of administering a brief paper-and-pencil task to them. The insulting was done by the author. A different experimenter obtained the TAT stories and ratings of subjective feelings, ostensibly as part of a study of "creativity." The situation was arranged in such a way that the subjects would not perceive a relationship between the aggression arousing situation and the subsequent measures.

The same procedure was followed for the other half of the subjects who served under Low Arousal conditions, with the exception that the paper-and-pencil task was administered in a friendly, nonprovoking way.

TAT Measure

Four TAT cards were used: 8BM, 2, 18GF, and 6GF. These TAT pictures, in the order named, were successively projected on a screen for 20 seconds. The subjects were given 4 minutes to write a story about each one. The usual group TAT instructions were used (McClelland, Atkinson, Clark, & Lowell, 1953).

The TAT cards were chosen to represent two levels of aggressive cue relevance. Cards 8BM and 18GF were assumed to be relatively high in aggressive cues (i.e., the pictures are easily interpreted as suggesting some aggressive action); Cards 2 and 6GF were assumed to be relatively low in aggressive cue relevance.

Each TAT story was scored in terms of the total number of aggressive acts mentioned in it.

Subjective Feeling Ratings

Immediately after completing their TAT stories, but before they were given any information regarding the true purpose of the experiment, all subjects were asked to rate the degree to which they had experienced certain feelings at the time they were about to begin their first TAT story. They were asked to estimate the extent they had felt "tired," "happy," "tense," "bored," "angry." For each of these feelings a subject's answer consisted of a check mark placed on a line 7 centimeters long ranging from Not at all to Very much so. The score for each item was the distance of the check mark from the Not at all end of the scale in centimeter units.

Maternal Questionnaire

When the subjects were told of the real nature of the study at the end of the experimental sessions they were also asked to supply the names and

addresses of their mothers. The mothers were subsequently contacted by mail and asked to complete a short, four-item questionnaire about their daughters. Two of these items are germane to the present report. They were designed to provide an index of the extent to which each subject's mother permitted the subject to aggress when the subject was a child. (Mothers were instructed to answer the questions in terms of how things were when the subject was less than 10 years old.) One item pertained to aggression toward parents, the other to aggression toward other children. Both were multiple-choice items containing four alternatives.[1]

A copy of the complete questionnaire used may be obtained from the author.

The responses to these items were combined by weighting each alternative in terms of the degree of permissiveness it indicated and adding the two scores. On the basis of a median split of these combined responses, each subject was categorized as having experienced as children either high maternal permissiveness (High MP subjects) or low maternal permissiveness (Low MP subjects) toward aggression.

The brevity of the maternal questionnaire apparently facilitated cooperation. Ninety-seven percent of the questionnaires sent out were completed and returned.

RESULTS

TAT Aggression

The mean aggression scores obtained for each of the eight cells of the design are presented in Table 1. A triple classification analysis of variance was performed on these data. The only significant main effect found was that of cue relevance of cards, with the cards defined as high in cue relevance for aggression eliciting significantly more aggression ($p < .001$) than those low in cue relevance.

TABLE 1

MEAN NUMBER OF AGGRESSIVE ACTS IN TAT STORIES

	High Arousal		Low Arousal	
	High MP subjects ($N = 23$)	Low MP subjects ($N = 27$)	High MP subjects ($N = 23$)	Low MP subjects ($N = 27$)
High aggressive cue cards	3.65_{b}	2.89_{bc}	2.74_{c}	2.85_{bc}
Low aggressive cue cards	1.87_{a}	1.89_{a}	1.78_{a}	1.22_{a}

Note.—Means with the same subscript are *not* significantly different from each other, by the Duncan (1955) multiple range test, at the .05 level.

The first prediction was not confirmed. The overall mean number of aggressive acts produced by High MP subjects (5.02) was greater than that produced by Low MP subjects (4.43), but this difference was not statistically significant ($p > .10$).

On the basis of Prediction 2 it was expected that there would be a significant interaction between the variables of permissiveness and arousal and/or between permissiveness and TAT aggressive cue relevance. Neither of these double interactions approached statistical significance. However the analysis of variance did reveal a significant triple interaction ($p < .02$) among the variables of permissiveness, arousal, and TAT cue relevance.

As can be seen in Table 1, Low and High MP subjects did differ in their reaction to aggression arousal in the direction anticipated in Prediction 2, if only responses to the High Cue cards are considered. High MP subjects when aroused showed a significant increase ($p < .05$) in aggression on the High Cue cards as compared with High MP subjects under Low Arousal conditions. Low MP subjects showed no such increase. In the case of responses to the TAT cards low in aggressive cues, however, aggression arousal had no differential effect upon High and Low MP subjects.

[1] These two items were:

1. Sometimes children get into quarrels with friends or other neighborhood children. To what extent did you permit your daughter to quarrel with children (other than brothers and sisters) when she was growing up? a. An effort was usually made to stop such quarreling immediately whenever it was observed. b. Some quarreling was permitted if it didn't last too long, or become a heated argument. c. Quarreling was usually permitted to continue, unless someone was really being hurt in the verbal battle. d. Such quarreling was almost never interfered with. It was considered to be a matter for the children to work out for themselves.

2. Sometimes young children become angry at their parents and say angry things to them. How much of this sort of anger did you permit in your daughter when she was growing up? a. None whatsoever at any time. b. It was *occasionally* all right for her to become angry at her parents if there was good reason but she was never permitted to be rude in expressing her anger. c. She was *usually* allowed to express her anger toward her parents when she felt angry, but open "sassing" of her parents was prevented whenever possible to do so. d. She was given full opportunity to express her anger, even when it meant being rude to her parents on occasion.

TABLE 2

MEAN SUBJECTIVE FEELING RATINGS

Feeling rated	High MP subjects ($N = 46$)	Low MP subjects ($N = 54$)	Significance of differences
Tired	4.41	3.80	*ns*
Happy	4.76	4.04	$p < .02$
Tense	2.59	3.57	$p < .02$
Bored	2.76	2.84	*ns*
Angry	1.12[a]	1.28	$p < .05$

[a] A square root transformation of ratings on the Angry dimension was performed after an inspection revealed unequal variances within cells.

Feeling Ratings

Separate analyses of variance were performed on the data obtained from each of the five feelings rated. Only in the case of the self-ratings on the Happy dimension did the results under High Arousal conditions differ from those obtained under Low Arousal conditions. Subjects under High Arousal conditions rated themselves as significantly less happy than control subjects ($p < .01$). In no case was there a significant interaction between the variables of arousal and degree of maternal permissiveness on these measures. However, on several of the feelings rated, High and Low MP subjects did differ in their mean self-ratings, irrespective of whether they served under the High or Low Arousal conditions.

Table 2 contains mean self-rating scores for High and Low MP subjects on each of the feelings rated. As this table indicates, Low MP subjects rated themselves as significantly less happy ($p < .02$), more tense ($p < .02$), and more angry ($p < .05$) than did the High MP subjects.

DISCUSSION

The model depicted in Figure 1 asserts that the net aggressive response strength of High and Low MP subjects should change differentially as arousal and/or cue conditions are increased. This principle was supported by the finding that insult produced an increase in the TAT aggression of High MP subjects, but not Low MP subjects. Moreover, this finding is consistent with the results of previous studies (Bandura & Walters 1959; Sears et al., 1957) which have indicated that maternal permissiveness toward aggression increases the likelihood of subsequent aggression.

Yet there is reason to doubt that the High MP subjects are necessarily more aggressive than the Low MP subjects under a variety of circumstances. Contrary to the expectation based on the model, they expressed no more TAT aggression in general than the Low MP subjects. Furthermore, the difference between High and Low MP subjects in reaction to insult occurred only on the TAT cards which had a strong "picture-pull" for aggression. This pattern of findings argues against a conception of the High MP subjects as individuals characterized by relatively strong, highly generalized readiness to aggress. Their heightened aggressiveness in the face of provocation and cues relevant to aggression can be viewed as an appropriate reaction to the stimulus situation. That the Low MP subjects showed no such increase in aggression, on the other hand, may be taken as an indication of inhibiting influences operating in these subjects. This interpretation proposes that High and Low MP subjects differ primarily in the extent to which they tend to inhibit the expression of aroused aggression rather than in the strength of their aggressive tendencies per se.

The subjects' ratings of their subjective feelings are noteworthy in this connection. The Low MP subjects, as compared with High MP subjects, rated themselves as feeling significantly more tense, more angry, and less happy. Their greater subjective tension and anger in the absence of even a comparable readiness to express overt aggression in the TAT situation supports the proposition that the Low MP subjects are more inclined to inhibit aggression than are High MP subjects.

Thus the data imply that relatively nonpermissive maternal attitudes toward childhood aggression are associated with the development of relatively strong internal restraints against expressing aggression. While the establishment of firm controls over aggressive tendencies may be regarded as a desirable goal, the self-ratings obtained in the present study also suggest that a more dysphoric subjective state may be another,

less desirable correlate of nonpermissive maternal attitudes.

SUMMARY

One hundred college girls told stories to two TAT cards high in aggressive cues and to two TAT cards low in cues for aggression. Half of the subjects told their stories after they had been subjected to an aggression arousing situation; the other half responded to the TAT cards without prior aggression arousal. The subjects were further categorized, on the basis of questionnaires completed by their mothers, as having experienced either relatively high maternal permissiveness (High MP subjects) or relatively low maternal permissiveness (Low MP subjects) toward aggression in childhood.

Two predictions were made: High MP subjects will express more TAT aggression in general than will Low MP subjects; High MP subjects will show a greater increase in TAT aggression as the cues and arousal conditions for aggression are increased than will Low MP subjects.

The data revealed that the High MP subjects did not express more TAT aggression in general than the Low MP subjects. High MP subjects did, however, show an increase in TAT aggression as arousal conditions were increased when responding to TAT cards high in aggressive cue relevance. No such increase occurred in the case of the Low MP subjects. Self-ratings made during the experiment indicated that the Low MP subjects felt significantly more angry, more tense, and less happy than the High MP subjects, irrespective of the degree of aggression arousal to which they were subjected.

REFERENCES

BANDURA, A., & WALTERS, R. H. *Adolescent aggression.* New York: Ronald, 1959.

CLARK, R. A. The effects of sexual motivation on phantasy. In D. C. McClelland (Ed.), *Studies in motivation.* New York: Appleton-Century-Crofts, 1955. Pp. 44–57.

DUNCAN, D. Multiple range and multiple *F* tests. *Biometrics,* 1955, **11,** 1–45.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.

SCOTT, W. A. The avoidance of threatening material in imaginative behavior. *J. abnorm. soc. Psychol.,* 1956, **52,** 338–346.

SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing.* Evanston: Row, Peterson, 1957.

WEATHERLEY, D. Anti-Semitism and the expression of fantasy aggression. *J. abnorm. soc. Psychol.,* 1961, **62,** 454–457.

(Received April 18, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 6–13

EXTERNALIZATION IN PERCEPTUAL DEFENSE [1]

DAVID W. LEWIT, A. RICHARD BRAYER, AND ALAN H. LEIMAN

University of Massachusetts

Perceptual defense has commonly been dealt with in experiments by determining recognition thresholds for disagreeable emotionally toned words and for neutral words, presented tachistoscopically. Various studies (see Goldiamond, 1958) have found higher thresholds for disagreeable words, when word frequency and other associative factors are controlled. This suggests that some mechanism of denial such as repression operates to inhibit recognition of such words, thereby reducing anxiety.

The possible role of anxiety in elevating thresholds is discussed by Allport (1955), Osgood (1957), and Eriksen (1958). It is clear that increased skin conductance, anxiety predispositions, previously established avoidance responses to stimulus words, and the rated disagreeableness of the words themselves are very often associated with elevated thresholds. But occasionally these factors are also associated with reduced thresholds, e.g., in Kurland (1954) and Spence (1957). This double effect suggests that under certain conditions defense mechanisms other than repression may operate to reduce anxiety and also to reduce thresholds for disagreeable words.

If someone outside the self were available, on which to project the motivation or responsibility for the word, then anxiety might be reduced as effectively as through repression. To the extent that projection could operate, defense by repression or denial would become less necessary or less functional. Thus we would expect that perceptual thresholds for disagreeable words for which an external source is provided might well be as low as thresholds for neutral words, and at any rate lower than for disagreeable words for which no external source is apparent.

To carry this reasoning one step further, it would seem that persons who generally deny responsibility—who tend to localize causation of everyday events outside themselves—would be influenced to a greater degree by an experimentally provided external source than would persons who tend to internalize causality. Thus we would apply chiefly to "externalizers" our assumption that perceptual thresholds for disagreeable words, when an external source is provided, will be about as low as thresholds for neutral words, and will be lower than thresholds for such words for which no external source is evident. Conversely, we would assume that "internalizers" would be less likely to take advantage of the external source provided, because in equivocal situations they more characteristically refer causality to themselves. Thus the internalizers' perceptual thresholds for disagreeable words would be higher than for neutral words regardless of the experimentally provided external source.

The present study tests these inferences concerning externalization, while controlling other conditions known to affect word recognition thresholds.

METHOD

Subjects and Personality Testing

The subjects were 134 male undergraduates enrolled in general psychology at the University of Massachusetts, 38 of whom served in Experiment I, and 96 in Experiment II. Immediately after the tachistoscopic procedure described below, each subject was asked to mark 25 items from the *Pd* scale of the MMPI (Hathaway & McKinley, 1951), 25 items from the Gough, McClosky, and Meehl (1952) scale of Social Responsibility derived chiefly from the MMPI, and a specially constructed scale of Externalization of Cause. The *Pd* and Social Responsibility items were chosen randomly, with the restriction that half of the items, if marked True, would contribute to a high score and half to a low score.

The Externalization scale appeared first in the test booklet bearing the title Situations Test. The instructions read:

> On the pages which follow are printed twenty statements about people and their reactions to

[1] Parts of this paper were presented at the meetings of the Eastern Psychological Association in New York on April 15, 1960. The research was in part supported by a grant of the Research Council of the University of Massachusetts.

various situations. After each statement there are three possible reactions to the situation. Write "S" in the space opposite the reaction which would be the *strongest* tendency in most normal people. Write "W" in the space opposite the reaction which would be the *weakest* tendency in most normal people.

Two typical items are:

Jack couldn't make it into one of the fraternities this year. He figures that . . .

—— he did not have the qualities that the fraternity was looking for
—— the fraternity had mistakenly heard he was more interested in another fraternity
—— he was probably blackballed because of the opinions of a few of the members who did not like him

Bob wins a new car on a TV quiz program. He figures that . . .

—— the quizmaster probably had orders to give the car away this time
—— his studying paid off
—— he was very lucky

Each item in the scale has one alternative which fixes the responsibility for the event upon the central character, another which fixes it upon someone else, and a third which attributes the cause of the event to an impersonal force. In the "Jack" example above these are the first, third, and second, alternatives, respectively, and in the "Bob" example they are the second, first, and third, respectively. The "Jack" example depicts an adverse event from Jack's point of view, while the "Bob" example involves a fortunate event. In the whole scale, 14 events were adverse and the remainder fortunate.[2] The subject's "externalization" score is the sum of the weights assigned to causes by forces other than the central character himself weighing S (strongest tendency) as 2, a blank as 1, and W (weakest tendency) as 0. Subjects with externalization scores above the median of the 134 scores were called "externalizers," and those with scores at or below the median were called "internalizers."

The *Pd* and Social Responsibility scales were included in the study for two reasons. Both of these scales and the Externalization scale should have in common a sensitivity to the "disowning" of responsibility, considering the Social Responsibility scale reversed, and might therefore produce parallel results with respect to perceptual defense. Also, they provide a basis for validation of the Externalization scale. In fact, the Externalization scale proved not to be significantly correlated with either of the other two scales, which themselves correlated −0.31. The reliabilities of the tests were 0.76 for the *Pd* scale, 0.77 for the Social Responsibility scale, and 0.51 for the Externalization scale, by retest over a 2-week interval. Despite the lack of correlation between the Externalization scale and the others, perceptual defense patterns involving *Pd* or Social Responsibility scores were similar to those reported below with respect to Externalization scores, though weaker and not statistically significant.

[2] Externalization scores based on only fortunate events bear a curvilinear relationship to those based on only adverse events. In the lower half of the range for fortunate events the relationship is negative; in the upper half it is positive or zero. For the 134 subjects of this study together with 168 others drawn from the same population, eta = 0.41 and $r = -0.28$. The hypothesis of rectilinearity is rejected by an epsilon prime square of 0.069, $p < .01$ with 8 and 293 *df*.

Stimulus Words

Table 1 shows the 20 common words selected as stimuli for this study. Five were emotionally neutral, 5 were associated with academic failure, 5 with physical violence, and 5 were obscene. The last three categories were chosen to correspond with achievement, destructive, and libidinal motives which are commonly viewed as basic for the culture from which the subjects were drawn. Thirty-five male undergraduates, independently drawn from the same population as the experimental subjects, rated 104 words for familiarity, frequency, emotional quality, and relevance to the above categories. After eliminating

TABLE 1

STIMULUS WORDS AND MEAN RATINGS OF THEIR CHARACTERISTICS

Stimulus words	Frequency	Familiarity	Disturbingness	Relatedness to		
				Failure	Violence	Obscenity
Neutral						
FIND	4.25	5.00	1.43	.23	.14	.20
LAKE	3.60	5.00	1.26	.00	.09	.21
BRING	4.83	5.00	1.63	.33	.33	.18
GARDEN	3.32	5.00	1.26	.17	.00	.26
STATION	3.60	5.00	1.80	.11	.23	.17
Mean	3.92	5.00	1.48	.17	.16	.20
Failure						
FAIL	4.83	5.00	3.83	3.86	.74	.23
POOR	4.11	5.00	2.92	2.47	.47	.44
CHEAT	4.23	5.00	3.20	2.77	.97	.71
STUPID	4.40	4.97	2.86	2.68	.40	.69
BLUNDER	3.03	4.92	2.63	3.29	.89	.35
Mean	4.12	4.98	3.09	3.01	.69	.48
Violence						
KILL	3.65	5.00	3.46	1.47	3.88	1.09
HATE	3.86	5.00	2.94	1.17	2.77	.51
CHOKE	3.43	4.97	3.00	1.06	3.37	.57
SCREAM	3.24	5.00	2.83	.80	3.23	1.14
BEATING	3.80	4.97	2.60	2.31	3.80	1.03
Mean	3.60	4.99	2.97	1.36	3.41	.87
Obscene						
PISS	3.74	5.00	2.68	.06	.12	2.38
CUNT	3.43	4.66	2.63	.44	.56	3.71
SHITS	4.08	5.00	2.88	.17	.26	2.77
FUCKED	4.14	5.00	2.40	.69	1.17	3.72
BITCHES	3.74	4.85	2.48	.57	.94	2.68
Mean	3.83	4.90	2.61	.39	.61	3.05

words of low or ambiguous relevance (as rated) to a category, five words for each category were selected for their match with words in each other category according to word length and these ratings of familiarity, frequency, and emotional quality.

Familiarity ratings were made on a six-point scale, for which "know exactly what the word means" was scored 5, "are almost sure of what it means" was scored 4, and "have never seen or heard the word" was scored 0. Distributions of ratings were J-shaped, and mean differences between the four categories proved insignificant by Friedman's analysis of variance by ranks (Walker & Lev, 1953). *Frequency* ratings were also made on a six-point scale. Raters were asked to rate the words "according to how often you see or hear them . . . in print or in conversation . . . in ordinary student life." "Very often occurs" was scored 5, "fairly often" 4, "occasionally" 3, . . . and "never" 0. The same test was used because of the highly skewed distributions, and intercategory differences proved insignificant. For *emotional quality*, raters were asked to "rate each word according to how pleasant or unpleasant it is to think about the word." Point values from 0 to 5 were assigned, respectively, to "very satisfying to you"; "somewhat pleasing"; "neither pleasing nor disturbing"; "slightly disturbing"; "quite disturbing, repulsive, or distressing"; "very disturbing, repulsive, or distressing to you." The similarity in quality of the failure, violence, and obscene word groups can be inferred from the failure of the Friedman analysis to find significant differences among them. In a separate analysis involving also the neutral word ratings, intercategory differences proved to be significant ($p < .001$).

The same students rated the *relevance* of each word to each of the three categories (or "topics") on five-point scales, from "very closely associated with the topic" to "completely *un*related to the topic." The differential relevance of the five failure words to the failure category, the five violence words to the violence category, and the five obscene words to the obscenity category is evident from the fact that 91% of these ratings, summed over words within categories, were in the "closely associated" half of the continuum, while 93% of other ratings were in the "unrelated" half.

Apparatus and General Procedure

Upon entering the dimly illuminated experimental room, each subject was asked to take a seat in relation to the screen as he would in a movie theater. Twenty seats were arranged in rows of four, with subjects seated from 12 to 24 feet from the glass beaded white screen. Between 4 and 17 subjects were tested at once. On the arm of each seat was a booklet with columns of blank lines for writing words to be shown on the screen. For practice the words VIEW, SEASON, FIVE, WATER, and BELIEVE—not to be used again—were tachistoscopically projected at various speeds and light intensities to accustom the subjects in 14 exposures to the constant position of the words on the screen, the progression from a very dim series of exposures through several stages to a very bright series, and recording after each flash.

Near the center of the screen were fixed two parallel horizontal yellow guide strips between which the words always appeared. One word at a time was projected from the rear of the room through an automatic projector using trays of prearranged slides and fitted with a shutter set at 0.04 second for all exposures after the preliminary training. Each word was typed in roman capitals on an electric typewriter, photographed, and mounted as a 35-millimeter slide to be projected to a height of 3.5 centimeters.

Thresholds for words were obtained by the ascending method of limits, modified so that all words were projected at a given light intensity before any word was projected at a higher intensity. The ambient illumination of the screen, and hence of the letters during any exposure, was 1.36 foot Lamberts. After the preliminary training and instructions to write "blank" if the subject could not see or guess the word, the 20 words were exposed in a random order at 6-second intervals. By changing Wratten neutral density filters mounted in front of the projector lens, successive exposure series had background illuminations of 1.50, 1.55, 1.61, 1.68, 1.79, 1.93, 2.11, 2.36, 2.67, 3.09, 3.64, and 4.38 foot Lamberts. For each of the first six series the words were presented in fresh random orders A, B, C, D, E, and F. The remaining six series presented these again in the rerandomized order D, C, B, E, F, A.

Any participant who recognized one or more words in the dimmest series or who failed to recognize any word by the brightest series was eliminated from the study. For any subject on any word, the threshold was taken as the ordinal number of the series on which the word was first accurately recorded.

EXPERIMENT I: PERSONALIZATION WITHOUT SET FOR CONTENT

Instructions

This experiment was designed to test the effect of personalizing the stimulus words—presenting them as having originated with a concrete person outside the self so that the subject could easily disown responsibility for them—while avoiding suggestion as to the content categories of the words. Therefore 16 subjects were instructed as follows, just before presenting the first series of exposures (Personalized-Nonset condition):

> Please turn to the next page—page 2 of your booklet. The words that will be flashed on the screen and that you will record on this page and the remaining pages of your booklet are the words of an actual person. In the spring of 1957 Dr. Kurt Blacker and his staff of the Department of Sociology at the University of California made tape recordings of over 300 spontaneous conversations in Los Angeles County as part of a research project to learn more about changing pat-

terns of everyday speech.[3] These candid conversations were picked up by concealed microphones in public buildings, vehicles, parks, and other locations. Recordings were made during various 24-hour periods over several months. Dr. Blacker has made available to us one of the conversations between two persons. The words which will be flashed for you in a moment were all taken from the remarks of one person in this conversation. The identity of this person was not made known to us.

During the brief rest periods after the first, fourth, seventh, and tenth series of exposures subjects were reminded briefly that the words were taken from this "speech study." Another group of 22 subjects serving as controls were told nothing of the origin of the words or the sort of words they might expect (Impersonal-Nonset condition).

Results

For each subject, the mean threshold for the five neutral words was subtracted from the mean threshold for the five words in the failure, violence, and obscenity categories, yielding three relative threshold scores. Relative threshold scores were used in order to avoid effects due to individual differences in such extraneous factors as visual acuity, distance from screen, and angle of viewing screen. The means of these relative thresholds are shown in Figure 1. The analysis of variance is shown in Table 2.

[3] The study and investigator alluded to here are fictitious.

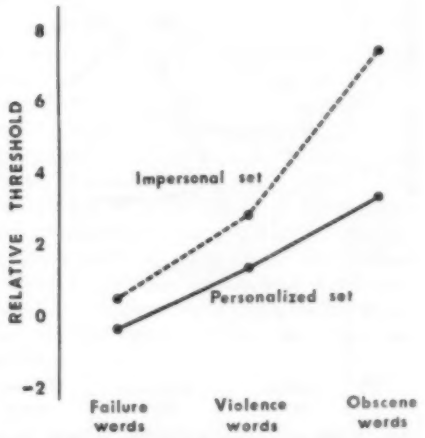

FIG. 1. Mean relative thresholds for word recognition under set for personal or impersonal source but in the absence of set for word content (Experiment I).

TABLE 2

ANALYSIS OF VARIANCE OF RELATIVE THRESHOLDS: EXPERIMENT I

Source	*df*	*MS*	*F*
Personalization (P)	1	17.59	6.40*
Externalization (E)	1	1.85	—
P × E	1	1.32	—
Error between subjects	34	2.75	
Word category (W)	2	28.45	12.73**
W × P	2	2.91	1.30
W × E	2	0.01	—
W × P × E	2	0.11	—
Error within subjects	68	2.23	

* $p < .05$.
** $p < .01$.

Two effects may readily be noted: thresholds for the three word categories are different, and thresholds in the Personalized-Nonset condition tend to be lower than in the Impersonal-Nonset condition. Both of these main effects are significant by analysis of variance, with $p < .01$ and $p < .05$, respectively. Externalizers showed much the same pattern of relative thresholds as internalizers, with no statistically significant differences distinguishing them in either main effects or interactions. Figure 1 therefore groups their data. No other interactions are significant. It may also be noted generally that without set for word category, most of the mean thresholds for disagreeable words are greater than for neutral words.[4]

EXPERIMENT II: PERSONALIZATION WITH SET FOR CONTENT

Instructions

In contemporary American culture stereotypes suggest that certain classes of persons are more likely to express libidinal urges, violence, or achievement striving than others. Ratings in the preliminary word study indicated that SAILOR, CRIMINAL, and STUDENT were at least moderately associated with "obscenity," "violence," and "failure," respectively. Persons of such classes, then, were chosen as appropriate specific targets for externalization where obscene, violent, and failure stimulus words are involved. The instructions attributing responsibility to a person of

[4] For Personalized-Nonset condition only, $t = 2.49$, $df = 15$, $p < .05$; for Impersonal-Nonset condition, $t = 5.45$, $df = 21$, $p < .01$.

one of these three classes for uttering the stimulus words begin with the same statements as those in the Personalized-Nonset condition, through the word "California." Instructions continue as follows:

Personalized-Obscene Set: . . . gathered information regarding the off-duty habits of naval personnel. As part of their study they recorded spontaneous conversations of members of ships' crews under various conditions. Dr. Blacker made available to us a tape recording of the voice of a sailor who had been ashore in Long Beach, California, for two hours with a shipmate. Both men worked in the engine room of the ship. The words which will be flashed for you in a moment were all taken from this recording. The conversation was picked up by a concealed microphone in a bar near the docks where sailors frequently go to drink and meet prostitutes. The remarks were recorded while the two men drank and talked in a booth near the door. They left the bar a half hour later with two women.

Personalized-Violence Set: . . . gathered information regarding the activities of criminal gangs. As a part of their study they recorded spontaneous conversations of gang members under various conditions. Dr. Blacker made available to us a tape recording of the voice of a young man who was a member of a criminal gang in Los Angeles. The words which will be flashed for you in a moment were all taken from this recording. The conversation was picked up by a concealed microphone in a detention room at the rear of a police station while he was alone in the room with another gang member. Both young men were arrested after a violent fight in which several people were severely injured.

Personalized-Failure Set: . . . gathered information on student adjustments to college. As a part of their study they recorded spontaneous conversations involving students under various conditions. Dr. Blacker made available to us a tape recording of the voice of a college freshman who had been called in for a talk with one of his professors. The student had been failing his course and not doing well in other courses. The conversation was picked up by a concealed microphone in the professor's office. The words which will be flashed for you in a moment were all taken from his remarks during this talk with his professor. The student left college at the end of the semester.

Instructions for corresponding Impersonal-Obscene Set, Impersonal-Violence Set, and Impersonal-Failure Set conditions were:

Please turn to the next page—page 2 of your booklet. The words that will be flashed on the screen and that you will record on this page and the remaining pages were taken from a list which included many words having to do with obscenities [violence] [failure]. Many of the rest of the words you will see are in this category, that is, words having to do with things dirty or obscene [fighting and violence] [doing badly in courses and failure].

For each personalized/impersonal pair of conditions set was controlled, allowing an interpretation of threshold variation in terms of externalization, unconfounded with simple associative factors. Between 14 and 19 subjects were used in each of the six conditions.

Results

Figure 2 shows mean relative thresholds for externalizers and internalizers separately. Means for failure words are based upon thresholds of subjects under failure set only. Means for violence words are based upon thresholds of subjects under violence set only. Means for obscene words are based upon thresholds of subjects under obscene set only. Table 3 shows the analysis of variance.

Mean relative thresholds for failure words tended to be negative, i.e., failure words were recognized more easily than neutral words, under failure set. Mean relative thresholds for

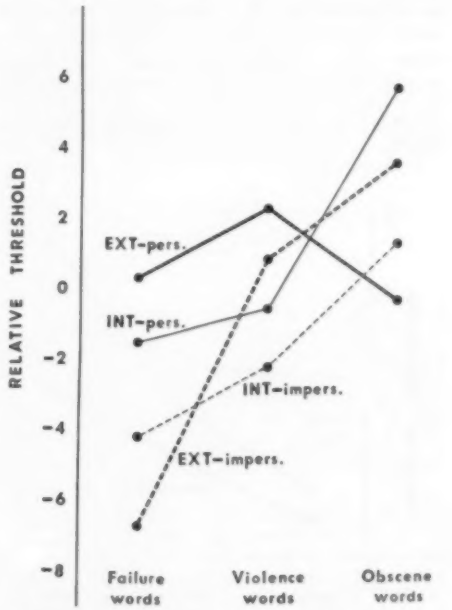

FIG. 2. Mean relative thresholds of externalizers (EXT) and internalizers (INT) for word recognition under personalized (pers.) and impersonal (impers.) set for source and word content (Experiment II).

TABLE 3

ANALYSIS OF VARIANCE OF RELATIVE THRESHOLDS: EXPERIMENT II

Source	*df*	*MS*	*F*
Word category (W)	2	31.92	10.03**
Personalization (P)	1	14.88	4.68*
Externalization (E)	1	0.14	—
W × P	2	5.27	1.66
W × E	2	6.75	2.12
P × E	1	1.43	—
W × P × E	2	10.00	3.14*
Error within subclasses	84	3.18	

* $p < .05$.
** $p < .01$.

violence words tended to be about zero, and for obscene words tended to be positive. This word category effect is significant at the .01 level in the analysis of variance. It can also be seen that the dashed lines are generally lower than the solid lines, indicating that impersonal set tended to reduce thresholds more than personalized sets. This effect is significant at the .05 level.

The interesting effect is the interaction of Set × Personalization × Externalization, which is also significant at the .05 level. This is due chiefly to the performance of externalizers given personalized sets. The subjects under these conditions show relative thresholds near zero for all stimulus words, while subjects under other conditions show considerable positive and negative deviations. On the basis of *t* tests, externalizers with an impersonal set showed mean thresholds for obscene and failure words which differ from zero at the .05 and .01 levels, respectively. Given a personalized set, however, none of their mean relative thresholds differ significantly from zero. Similar tests on means of internalizers show no significant deviations from zero under impersonal set, but under personalized set show a deviation for obscene words only ($p < .02$). Neither the externalization main effect nor any two-factor interaction was significant.

DISCUSSION

The central hypothesis of this study was that projection will operate to reduce recognition thresholds for disagreeable words provided that (*a*) the person concerned is predisposed to externalize responsibility, and (*b*) a target is available upon which to project responsibility.

The results of Experiment I confirm the hypothesis if Provision *a* is disregarded—when subjects believe that the stimulus words were originated by an anonymous and randomly selected person who spoke on an unspecified topic without knowledge that the utterances were being overheard, they tend to recognize several kinds of disagreeable words more readily than when they have no such basis for attributing cause to a person outside the experimental situation. It did not seem to matter, however, whether the subject was predisposed to externalize cause.

A more parsimonious explanation—response bias—is plausible for Experiment I. It would hold that a personalized set would tend to restrict responses to groups of thematically related words since they presumably originated in a single conversation. Once any word is recognized, a related word should be a more probable response to any subsequent stimulus. An analysis of responses following the first correct response lends no support to this supposition—neither the next word guessed nor the next correctly named word fell any more frequently in the same word category among subjects given the personalized instructions than among those given impersonal instructions. The latter do, however, show somewhat more perseveration from the five practice words before making a correct identification, which may interfere with accuracy of response. The results of Experiment II, however, show generally higher rather than lower thresholds for subjects under personalized conditions. In any event, instruction induced response bias cannot be an adequate explanation for the externalization interaction effect of Experiment II, apart from personalization per se, since experimental conditions for externalizers were identical with those for internalizers.

Experiment II provided several targets, each more specific than that in Experiment I. Since the topic of the utterances of these targets was clearly implied in the instructions, it became necessary to induce an impersonal set for the same categories in subjects for whom no target was to be made available. In

this experiment the hypothesis including both Provisions *a* and *b* was confirmed, but apparently only for obscence words. While results for violence words revealed nothing, results for failure words were roughly the opposite of those for obscene words, i.e., externalizers who had a target showed the highest thresholds to the failure words, and those without one, the lowest.

But is the failure word effect opposite to the obscene? An implicit assumption in the theoretical discussion early in this paper is that generally, thresholds for any sort of disagreeable word would be higher than for neutral words—threshold reducing conditions are postulated only for positive relative thresholds. Actually failure words were generally recognized sooner than neutral words, i.e., negative relative thresholds were involved. If thresholds either decidedly higher *or* decidedly lower than for frequency matched neutral words are taken as indicators of some defensive process, i.e., repression or vigilance, then it can be said that externalizers with targets were not exhibiting such a defense with respect to either obscene or failure words. Since without targets externalizers did exhibit defensive behavior just as internalizers did, then the offering of a target seems to induce a process of "secondary defense" or "perceptual counterdefense" in externalizers. With respect to the hypothesis stated above, both Provisions *a* and *b* hold if we concern ourselves with the *shifting* of relative thresholds toward zero (neutral word level) rather than with threshold reduction as such.

A discrepancy exists between Experiments I and II when we consider that internalizers as well as externalizers responded to the personalization procedure in Experiment I. One important difference in the conditions of these two experiments besides set for a certain category of word should be noted. The "sailor," "student," or "criminal gang member" of Experiment II was a clearly described, recognizable target whereas the person overheard in the other experiment's "speech study" was nonspecific and not identifiable. If we assume that internalizers are generally less prejudiced persons than externalizers, then they are likely to have felt guilt whenever they blamed others for acts resulting from the other's poor condition. Under the circumstances of Experiment II, therefore, any tendency to externalize may have been inhibited. Under the diffuse personal specifications of Experiment I, however, such a tendency could be expressed without discrimination against any class of person and therefore without anxiety. Thus relative thresholds of internalizers may approach zero (neutral word level), as do those of externalizers, when a nonspecific target person is made available.

In light of these considerations, then, the central hypothesis might be restated for further investigation as follows: projection will operate to shift relative thresholds for disagreeable words toward zero (neutral word level) provided that (*a*) the person concerned is predisposed to externalize responsibility, (*b*) a target is available upon which to project responsibility, and (*c*) the specific identity of the target does not arouse guilt in the perceiver.

Summary

Perceptual defense studies have usually found high recognition thresholds for disagreeable words. Sometimes, however, thresholds have been reported which are the same as those or lower than for neutral words. It is postulated that defense mechanisms other than repression may account for the low thresholds. To test projection as a perceptual defense mechanism, subjects who tended to externalize cause were identified and an opportunity was provided to project onto others the responsibility for disagreeable words.

Threshold reduction for obscene, violence, and failure related words was obtained for both externalizers and internalizers, when an unidentified human source for the words was supplied. Neutralization of both high and low thresholds for disagreeable words was obtained for externalizers but not for internalizers, when more clearly identified human sources were supplied.

It is concluded that sensitization and inhibition may both be primary perceptual defenses, but that secondary or counterdefenses may operate to neutralize these effects when conditions for the externalizing of responsibility obtain.

REFERENCES

ALLPORT, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.

ERIKSEN, C. W. Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958.* Lincoln: Univer. Nebraska Press, 1958. Pp. 169–227.

GOLDIAMOND, I. Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychol. Bull.,* 1958, **55,** 373–411.

GOUGH, H. G., MCCLOSKY, H., & MEEHL, P. E. A personality scale for social responsibility. *J. abnorm. soc. Psychol.,* 1952, **47,** 73–80.

HATHAWAY, S. R., & MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory.* (Rev. ed.) New York: Psychological Corporation, 1951.

KURLAND, S. H. The lack of generality in defense mechanisms as indicated in auditory perception. *J. abnorm. soc. Psychol.,* 1954, **49,** 173–177.

OSGOOD, C. E. Motivational dynamics of language behavior. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957.* Lincoln: Univer. Nebraska Press, 1957. Pp. 348–424.

SPENCE, D. P. A new look at vigilance and defense. *J. abnorm. soc. Psychol.,* 1957, **54,** 103–108.

WALKER, HELEN M., & LEV, J. *Statistical inference.* New York: Holt, 1953.

(Received April 18, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 14–25

DREAM REPORTS FROM DIFFERENT STAGES OF SLEEP[1]

W. DAVID FOULKES[2]

University of Chicago

The recent discovery of objective techniques of dream detection (Dement & Kleitman, 1957b) suggests the possibility that dream formation can be observed directly rather than inferentially. The present study is an attempt to explore this possibility in light of recent studies of EEG and eye movement correlates of dreaming.

Dement and Kleitman (1957a) have described four stages of EEG activity during sleep. Stage 1 is characterized by the complete lack of spindles and by low voltage nonperiodic activity with occasional bursts of alpha; Stage 2 is characterized by spindling against a low voltage background; Stages 3 and 4 by slow high voltage activity; Stage 3 also has spindles while none occur in Stage 4. These latter two stages are "deep" sleep, while Stage 1 is the "lightest" stage of sleep. EEG activity shows a cyclical nature during sleep. A rapid progression from Stages 1 to 4 occurs after sleep onset, with Stage 4 persisting for varying amounts of time, then yielding, in turn, to Stages 3, 2, and 1. Consistently and exclusively associated with this emergent[3] Stage 1 activity are bursts of rapid eye movements (REMs). When the first rapid eye movement period (REMP) ends, the EEG progresses from Stage 1 to Stages 3 or 4 as the cycle recurs. Later repetitions of the cycle differ from the first cycle in that the probability of reaching Stage 4 between REMPs drops sharply after the second cycle, and in later cycles REMPs become longer and the time between them shorter.

As a result of a study of a number of subjects, Dement and Kleitman (1957b) noted that dream reports were obtained almost exclusively at awakenings preceded by Stage 1 with concurrent REMs and were obtained almost always under these conditions. These investigators concluded that Stage 1 may be the better index of dreaming since dreams may lack visual imagery. Other publications from this group of investigators at the University of Chicago have tended to confirm the assignment of dreaming to Stage 1 REMPs, while also elaborating on the above findings in various respects (Dement, 1955; Dement & Wolpert, 1958a, 1958b; Wolpert, 1960; Wolpert & Trosman, 1958). Goodenough, Shapiro, Holden, and Steinschriber (1959) have independently arrived at similar conclusions about the location of dreaming in the sleep cycle, although it should be noted that their evidence is much more equivocal on the proposition that dreams occur *only* during REMPs.

The apparent consistency of these recent findings that dreaming is associated with a particular stage of ocular activity (REMPs) and with a particular stage of EEG activity (Stage 1) and that dreaming occurs only intermittently during the night suggested the desirability of trying to determine more precisely the point of dream onset in terms of EEG and REM activity. Experimental awakenings made at various intervals after the occurrence of the objective index of dream onset might provide an empirical basis for theorization about the usual sequence of dream development.

A pilot study was undertaken to gather preliminary data concerning the exact point at which dream onset might be said to occur. In addition, information was sought about the phenomenological nature of the onset of dreams and of temporal trends following such

[1] This paper is based upon a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Chicago. The research was supported by the National Institute of Mental Health, United States Public Health Service, under Research Grant M-2116 (C-2) to Joe Kamiya, whose encouragement and assistance in the conduct of the research are gratefully acknowledged. Helpful suggestions were also received from Donald W. Fiske and Austin H. Riesen. The author held a National Science Foundation Cooperative Graduate Fellowship during this investigation.

[2] Now at Lawrence College.

[3] Emergent refers to the later phase of Stage 1 in which sleep lightens.

onset. EEG and REM recording procedures were almost identical with those employed by Dement and Kleitman (1957b). Three subjects were run for a total of 25 nights.

The results of the pilot study are reported in greater detail elsewhere (Foulkes, 1960). Briefly, the results indicated that recall was most likely following control REMP awakenings made well after REM onset (96% mean subject value), but that recall was also elicited quite often following experimental awakenings made within a minute or two of Stage 1 onset, apparently regardless of whether or not REM onset had occurred (67% mean subject value). Although recall figures from outside Stage 1 were considerably lower (34% mean subject value) than those within it, they were considerably higher than those reported in previous Chicago studies, most consistently so for pre-REMP Stage 2 awakenings (49% mean subject value). For both deep sleep awakenings (Stages 3 and 4) and Stage 2 awakenings, there was almost the same frequency of recall at those awakenings without a prior REMP for which the subject was not awakened and had not made a report as there was for awakenings with a prior undisturbed REMP. This result suggested that the recall obtained outside REMPs could not be explained as representing delayed recall of earlier, unreported REMP dreams, an explanation offered by Dement and Kleitman (1957b) to account for the few cases of dream recall they found outside REMPs.

The finding that recall could be elicited earlier in the sleep cycle than reported in most of the prior studies led to an examination of the qualitative differences of reports obtained at various points leading into and during early intervals of Stage 1 and REM activity. An impressionistic analysis of reports from pre-REMP Stage 2, Stage 1 before REM onset, and Stage 1 at and immediately following REM onset suggested the operation of a developmental process over this portion of the sleep cycle involving a shift from the conceptual to the perceptual mode of experience, a shift from fragmentary and disjointed content to elaborated and integrated content, and increasing distortion and decreasing reality-contact of content.

Although the qualitative findings indicated that the methods used in the pilot study might contribute to our understanding of dream formation, the recall findings cast doubt upon the original assumptions of the study. Dream reports were elicited with some success in all categories of awakening. Dreaming, or some form of cognitive and perceptual activity, thus might conceivably occur continuously in the sleeping human and REMPs might be useful only in differentiating different kinds of dreams, rather than as signs of dreaming versus nondreaming.

On the other hand, even if dreaming does occur all night, we may still ask how those distinctive (hence often noticed) dreams which occur in Stage 1 REMPs seem to "originate," for the pilot study results definitely suggest that there is a developmental process occurring just prior to and during the early parts of these periods. In light of the pilot study results, however, it seemed desirable to broaden the focus of subsequent research to include possible qualitative differences in mental phenomena during sleep, whenever they might occur.

METHOD

Procedures

A four channel Offner electroencephalograph was used to record brain potentials. The following leads were employed: parietal lobe versus both ear lobes; occipital lobe versus both ear lobes; and one eye referred to the other. The eye electrodes were placed lateral to the eyes. Continuous EEG recordings were made from the time the subject went to bed until the time at which he wished to arise in the morning, with the exception of the time spent in reporting during awakenings. Technical procedures closely paralleled those used by Dement and Kleitman (1957b).

Experimental awakenings were made by a loud bell, except for Subject 102 who was awakened by the experimenter's voice. Immediately after the bell was rung, the experimenter, via an intercom system, would call the subject by name to make sure that he was awake, and then would ask him if he were dreaming. If the subject indicated that he was not dreaming, he was then asked if anything was "going through his mind." This latter question was used as it was suggested by the pilot study and confirmed in this study that not all cognitive and perceptual phenomena occurring during sleep are labeled by subjects as "dreams." The subject was then encouraged to report his "dream" or "thought" or whatever it was that he felt to be "going through his mind" before the bell rang. When he indicated

that he was finished, he was asked a number of standardized questions about the dream content. The questions were as follows:

What was the setting of this dream?
Was the setting at all familiar?
What other people were present in this dream?
Were any of them familiar?
Were you yourself present in this dream?
Did you have any feelings or emotions in this dream?
Were others experiencing feelings or emotions? (Not always asked)
Were you dreaming at the time the bell rang? (If no) When did this dream take place? (If yes) What was happening just as the bell rang?
Was this dream at all visual? (If yes) What scenes do you recall?
Was there any conversation in this dream? (If yes) Were you yourself talking?
Did the scene shift at all in this dream?
Were you aware that you were dreaming?
Was the imagery clear in this dream?
Did the content of the dream seem coherent to you at the time of its occurrence?
Was the dream in color? (If yes) Is there anything in particular that you recall as having been in color?
For about how long do you think that you had been dreaming when I rang the bell?
Do you think that there were any other dreams or that anything else was going through your mind this time, even if you cannot recall what they or it might be?
Did you engage in any physical activity in this dream? Did others?

When reports were not visual and did not seem to have any of the characteristics of what we usually call dreams, a number of these questions were inapplicable.

To ensure that subjects, faced with the somewhat unpleasant experience of successive awakenings in the middle of the night, would not attempt to return to sleep immediately by grunting out a few drowsy, one-syllable replies, a continued arousal technique was used. Subjects faced a reporting session of several minutes length, even if initially upon awakening they definitely felt that nothing had been going through their mind prior to the awakening bell, or did not seem to be able to recall anything when claims of dreaming or thinking were made. The experimenter did not attempt, during this continued arousal period, to keep the subject awake and alert by suggesting that he was indeed dreaming on occasions when he had previously claimed otherwise. The questions used in these instances suggested that the experimenter was satisfied that the subject had experienced no conscious phenomena prior to the awakening, but was now interested in the character of his sleep.

Upon arising in the morning the subject was asked to try to recall all the dreams that he had related during the night. Since the subject was, after this recall testing, to fill out a rating form for each of his reported dreams, it was felt important to ascertain that he was doing this on the basis of as close to the same information that he had at his disposal right after their occurrence as possible. The overwhelming majority of reports made during the night were either freely recalled or completely recalled after a prompting from the experimenter about one element therein.

Most of the questions on the rating form, the Dream Information Survey, paralleled and supplemented—in six-choice rating scale form—the qualitative information probes made during the actual reporting of the dream. Subjects were asked to rate the dream on the following dimensions: amount of activity; degree of control experienced over events in the dream; emotion experienced, self; emotion experienced, others; felt anxiety; clarity of content; distortion; relation of content to own personal life; pleasantness-unpleasantness; frightening quality; dramatic quality; activity-passivity of role of self; amount of violence or hostility in the dream; and degree to which they make sense of the dream. In addition, subjects were asked: to say how many different scenes or distinct parts there were in the dream, to list them, and to check those which seemed to have had clear visual imagery; if they found themselves in the dream doing anything which for ethical reasons they would not do in real life, and if so, what; and, if the dream was scored as making some sense to them, to tell how it made sense to them. These questions, in conjunction with those asked at the time of the original report, were designed to provide qualitative and quantitative data with which to compare dreams obtained from different conditions of sleep and ocular activity. As was the case with the questions asked during the original report, some of the Dream Information Survey questions proved inapplicable to reports which the subject classified as "thoughts." For such reports, the only responses tabulated were ratings for two relevant dimensions (relation of content to own personal life and degree to which sense is made of the dream) and responses to the open-end question as to how sense was made of what had been reported.

Experimental nights were generally scheduled once a week. Each subject was scheduled to serve for 7 nights. Due to an inability to return to sleep after his first awakening on his first night, one subject (109) was run for 8 nights.

The following categories of awakening were used:

1. Long REMP (four awakenings, one each at 9, 12, 18, and 24 minutes after the onset of REMs).
2. Medium REMP (four awakenings, one each at 2, 3, 4, and 6 minutes after the onset of REMs).
3. Brief REMP (four awakenings, one each at 60, 30, and 15 seconds after REM onset, and one at as close to REM onset as the experimenter's judgment and reaction time permitted, generally about 3 or 4 seconds).
4. Stage 1 prior to any REMs (four awakenings).

5. Stage 2, made at intervals sufficiently long since the onset of sleep or prior REMP to lead the experimenter to believe that Stage 2 would shortly issue into Stage 1 (four awakenings).

6. Hypnagogic awakenings, made following either the initial onset of spindles or spindle onset after an awakening, mostly in Stage 2 (four awakenings).

7. Stages 3 and 4 (four awakenings).

Four awakenings were scheduled on each of the 7 nights. The schedule controlled for the effects on recall incidence of position of awakenings during the night and within the series of nights. It was not always possible to follow this schedule exactly, but a reasonable approximation to it was achieved.

Subjects

Eight subjects were run according to the procedures outlined above. They were all paid volunteers, males, and ranged in age from 17 to 27.

Subjects were interviewed prior to their first night. They were told that the experiment was concerned with learning certain facts about the process of dreaming, rather than facts about them as persons. They were informed that the experimenter would not necessarily know whether or what they were dreaming at the experimental awakenings; this, rather, was what he meant to investigate. The first experimental night was never scheduled on the day of the interview.

Classification of Reports

It became evident early in the present research that reports tended to be classified by the subjects and to be readily classifiable by the experimenter as falling into two distinct classes: some reports were of "dreams," while the others were of "thoughts." The following criteria (essentially those used by the subjects themselves) were used in classifying reports with some substantive content as dreams or thoughts: any occurrences with visual, auditory, or kinesthetic imagery were called dreams; phenomena reported to lack such imagery but in which the subject either assumed another identity than his own or felt that he was thinking in a physical setting other than that in which he actually was (i.e., the laboratory) were also classified as dreams; phenomena lacking sensory imagery (no imagery other than visual, auditory, or kinesthetic was ever noted) and changed identity or setting were classified as thoughts. A typical thought report is the following:

> I was thinking of tax exemption, and it skipped around to something else. . . . [Experimenter: What were you thinking about tax exemption?] Oh, one of the points made in the class . . . was that you had to provide over half of the support of a person in order to claim them as dependents.

It should be noted that the criteria used to differentiate dreams from thoughts are primarily formal in character. They are concerned, except in a few borderline cases, with the presence or absence of sensory imagery, not with the degree of correspondence of the reported material to the subject's waking life. Thus an accurate recollection of some prior experience in the subject's life could be classified as either a dream or a thought, depending on whether or not it had sensory imagery.

To be classified as producing thinking or dreaming content, a report would have to include some item of specific content. Excluded from the analysis of content was the following material: material reported after a subject had napped briefly while talking with the experimenter (to exclude material which might be hypnagogic reveries from lapses into sleep during reporting); new material presented during the morning recall testing session; material from awakenings where the objective record was unclear or did not satisfy the criteria for the category intended; and dreams or thoughts felt by the subject to represent material which occurred considerably prior to the awakening bell. With regard to this last class of material, it should be noted that reports of dreams which the subject felt had occurred considerably prior to the bell (there were seven such reports) were obtained only in conjunction with other reports which were localized at the ringing of the bell, and thus the exclusion of such dreams did not affect the recall incidence figures.

Although previous investigators have reported wholly unsubstantiated claims of dreaming as "not dreaming," it seemed desirable to distinguish between those times when no specific content was produced but claims of dreaming or thinking were made and those times when the subject felt quite definitely that nothing had been "going through his mind" just prior to the awakening bell. In the present analysis, therefore, the following five classifications of recall are used: dreaming content, thinking content, dreaming claim, thinking claim, neither dreaming nor thinking.

RESULTS

Incidence of Recall

Table 1 presents the pooled recall results for REMP (Categories 1 through 3) and non-REMP (Categories 4 through 7) awakenings. The column labeled Consistency gives information on the degree to which the pooled data reflect the data of the subjects considered singly. The first entry gives the number of subjects who show a difference between REMP and non-REMP proportions with a sign identical to that in the pooled data; the second figure gives the number of subjects who show a difference in proportions with a different sign than that in the pooled data; and the third figure gives the number of subjects whose data show identical proportions in the REMP and non-REMP summation of

categories. All REMP versus non-REMP differences in this and subsequent tables were tested by the Wilcoxon matched-pairs signed-ranks test (Siegel, 1956). Each subject contributed a pair of observations, a proportion in REMPs, and a proportion outside such periods. All tests were two-tailed.

Table 1 reveals that recall of some item of specific content was more likely to occur in awakenings during REMPs than in awakenings outside REMPs. This finding is consistent with the results of all prior studies. The high frequency of recall outside REMPs is at variance with the results of all recent systematic attempts to relate dreaming to EEG activity, with the possible exception of that

TABLE 1

RECALL:
DURING AND OUTSIDE REMPS

Recall category	REMP		Non-REMP		Consistency
	n	Proportion	*n*	Proportion	
Dreaming content	89	.82	73	.54	8-0-0**
Thinking content	5	.05	27	.20	6-0-2*
Content total	94	.87	100	.74	6-1-1*
Dreaming claim	3	.03	8	.06	—
Thinking claim	2	.02	10	.07	—
Content plus claim total	99	.92	118	.87	5-2-1
Neither dreaming nor thinking	9	.08	18	.13	—
Total	108	1.00	136	1.00	

* Significant at .05 level.
** Significant at .01 level.

of Goodenough et al. (1959). It should be noted that all eight subjects had a higher frequency of recall of dreams in REMP awakenings than in non-REMP awakenings, and that thinking was more likely to occur outside REMPs than it was during them. It should also be noted that the gap between "content plus claim" frequencies for REMP and non-REMP awakenings is quite small.

Since the data of Table 1 are pooled over the initial categories of awakening, it might be asked if any one non-REMP category contributed a disproportionate share of the recall in the non-REMP classification. The data in Table 2 show that this is not the case.

For further clarification of the data in Table 1, Table 3 presents the individual subject dreaming and thinking recall frequencies for REMP and non-REMP awakenings. Data on claims are not included as their frequency was so slight, but it should be noted that Subject 102, whose level of non-REMP recall was considerably lower than the over-all average, contributed 9 of the 18 non-REMP claims, 8 of them thought claims. The mean proportion of recall for all subjects was .74 in non-REMP awakenings and .87 in REMP awakenings. No subject failed to recall at

TABLE 2

RECALL:
BY CATEGORY OF AWAKENING

Awakening category	Content			Content plus claims total	*n*
	Thinking	Dreaming	Total		
Hypnagogic	.25	.56	.81	.91	32
Stage 3/4	.19	.51	.70	.86	37
Stage 2	.23	.51	.74	.86	35
Pre-REMP Stage 1	.12	.56	.69	.84	32
Brief REMP	.03	.83	.86	.86	36
Medium REMP	.06	.79	.85	.91	34
Long REMP	.05	.84	.89	.97	38

TABLE 3

RECALL:
BY SUBJECT

Subject	Recall	REMP		Non-REMP	
		n	Proportion	*n*	Proportion
102	Dreaming content	9	.69	1	.06
	Thinking content	2	.15	7	.39
	Total content	11	.85	8	.44
104	Dreaming content	13	1.00	15	.88
	Thinking content	0	.00	2	.12
	Total content	13	1.00	17	1.00
105	Dreaming content	11	.85	13	.81
	Thinking content	0	.00	1	.06
	Total content	11	.85	14	.88
106	Dreaming content	9	.75	9	.53
	Thinking content	1	.08	3	.18
	Total content	10	.83	12	.71
107	Dreaming content	15	1.00	11	.69
	Thinking content	0	.00	0	.00
	Total content	15	1.00	11	.69
108	Dreaming content	11	.79	12	.71
	Thinking content	0	.00	0	.00
	Total content	11	.79	12	.71
109	Dreaming content	14	1.00	10	.56
	Thinking content	0	.00	7	.39
	Total content	14	1.00	17	.94
110	Dreaming content	7	.50	2	.12
	Thinking content	2	.14	7	.41
	Total content	9	.64	9	.53

least one dream or thought in each of the seven awakening categories.

Table 3 reveals that 25 of the 32 thinking reports were contributed by three subjects, 102, 109, and 110. Two of these subjects, 102 and 110, had the lowest over-all recall rates, and this suggests that thinking reports may represent the efforts of subjects with "elusive" dreams to catch hold of phenomena which could well be dreams, but whose nonperceptual content is all that can be grasped. Subject 110, however, had the highest score of any subject on the Primary Mental Abilities space factor, a measure of facility in visualization, and this suggests that his thinking reports may represent a real phenomenon and not an artifact of incomplete reporting of essentially visual phenomena.

Subject 110's recall figures are most interesting in view of his motives for volunteering his services. He had heard of studies being conducted in the area of dream research, and, never aware of having dreamed, he was desirous of seeing whether he would be an exception to the statements he had heard to the effect that everyone dreams. This subject, whose sensitivity to the processes of his sleeping consciousness was evidently very low, provided a stiff test for the proposition that all subjects could report dreams from any stage of sleep or ocular activity. As Table 3 indicates, this subject reported specific content only 17% less often outside REMPs than he did within them, though the absolute level of recall in both conditions was lower than average.

The recall results appear to offer some support for our earlier suggestion that dreaming may be present throughout the sleep cycle. If our findings from awakenings made around the time of REM onset prove to be at all relevant to the study of dream formation, then, they will be relevant to the formation of a particular class of dreams, not of all dreams—if by dreams we mean any mental activity during sleep. While the initial choice of awakening categories served to concentrate attention on awakenings close to presumed REM onset or soon after REM onset, the procedures of investigation employed should, in view of the recall data, aid us in answering a more general question concerning qualitative differences among reports obtained from the whole range of conditions of the human sleep cycle.

Qualitative Characteristics of the Reported Phenomena

Information on the qualitative characteristics of the dreams and thoughts reported by the subjects, as well as on certain other characteristics, came from two sources, the responses subjects made to probes during the original dream report and their responses to the Dream Information Survey.

Original Report. Table 4 indicates those questions asked during the subject's original report which showed significant differences in response between REMP and non-REMP

TABLE 4

SIGNIFICANT DIFFERENCES IN QUALITATIVE CHARACTERISTICS OF REPORTED PHENOMENA[a]

Dimension	REMP		Non-REMP		Consistency
	n	Proportion yes	*n*	Proportion yes	
Formal differences: Question					
Emotion self?	88	.50	95	.32	7-0-1**
Emotion self or others?	92	.55	96	.31	8-0-0***
Visual?	93	.90	100	.66	7-0-1**
Visual imagery: clear?	79	.80	61	.62	6-1-0*
Physical movement: self?	73	.67	57	.38	7-1-0*
Locomotion: self?	73	.42	57	.21	8-0-0***
Locomotion: others?	66	.45	51	.25	6-2-0*
Only one other character?	79	.34	60	.57	7-0-1**
More than one part?	93	.37	99	.17	6-0-2*
Scene shifts?	81	.63	60	.33	7-0-0**
Median judged length (dreams only)	82	(5m)	64	(3m)	6-1-1*
Correspondence differences: Dimension					
Manifest work/school theme	94	.15	100	.27	7-1-0*
Continuation of prior dream/thought	62	.11	75	.33	7-1-0**
Memory processes	94	.01	100	.17	7-0-0**

[a] Source: Subjects' original reports.
* Significant at .05 level.
** Significant at .02 level.
*** Significant at .01 level.

awakenings. The differences have been put into two classes, according to whether they deal primarily with formal characteristics of the reported phenomena or with the correspondence between dream and thought content and prior waking experience of the subject. Except where indicated, the *n* refers to both dreaming and thinking reports where the question was one which was asked with both kinds of reports. The number of cases analyzed in certain categories is often less than the total number of appropriate cases. Subjects sometimes found it impossible to reply either affirmatively or negatively to the probe, and all probes were not always asked, due to oversight on the experimenter's part.

The important formal differences seem to fit into two general areas: organismic involvement and elaboration. With regard to the former, the REMP reports more often contain affective involvement on the dreamer's part (and this difference is independent of whether thinking reports are excluded from or included in the analysis). The greater incidence of involvement of the visual system also strikingly differentiates the reports obtained in and outside REMPs. A further index of visual involvement shows that visual imagery is not only more often present in REMP reports, but also, when present, it is more often "strong" in quality. Physical movement, especially gross movement (locomotion), also is found more often in REMP dreams than in non-REMP dreams. It should be noted that many of these indices successfully differentiate REMP and non-REMP *dreams*, since many of the questions apply only to such reports. As we have seen, that form of mental activity with the least organismic involvement, ***thinking***, occurred almost exclusively outside REMPs.

REMP dreams more often involved other people (this difference was not significant), and when others were present, it was more often more than one person. REMP reports more often contained more than one distinct part. REMP dreams were generally felt to be longer, and more often involved scene shifts. These features connote a degree of integration and continuity in REMP dreams greater than that in non-REMP dreams.

Because of the theoretical and interpretative significance of responses to the question as to whether the reported material occurred at the time the bell rang, it should be noted that thoughts were always referred to the period of sleep immediately prior to the awakening bell, and that dreams from outside REMPs were very often (88% of cases), though less often than REMP dreams (98% of cases), referred to this period just prior to the bell. In fact, Stage 2 dream reports were felt to represent material occurring as the bell rang as often as were reports from the longest REMPs. Even when cases where the subjects claimed they were dreaming as the bell rang but could not recall what was happening just at that point were excluded, the REMP and non-REMP proportions of affirmative response were almost completely unchanged.

All subjects were either employed and/or active students during their experimental service. The largest single block of their waking life thus usually consisted of job performance or studying. If dream and thought content correspondence to waking life were to be assessed, this seemed an obvious place to start. To be scored affirmatively, the dream or thought had to involve directly the subject's work or study element in waking life. Subjects' comments on the Dream Information Survey as to how they made sense of the dream were utilized in scoring this dimension in addition to the subjects' comments at the time of the original report.

The most obvious recent waking event for dreams or thoughts reported after the initial recall producing awakening of a given night would be the report of the preceding dream or thought on that night. Although some sort of link could probably be traced from one dream or thought to the next in all cases, the present analysis limits itself to those dreams or thoughts which seemed to continue the larger aspects of the prior report—the topic or situation of the last reported thought or dream.

Subjects sometimes stated during the night that what they had reported was either a recreation of a past event, thought, or situation in their experience or that it was a thinking back to some past event, thought, or situation which left that past element completely

undistorted, i.e., as it actually had been in their experience. An example of such a report is the following:

> I was thinking about this phone call I got from home tonight. I got a long-distance call from home, and my parents were very angry with me because I haven't written. . . . They really gave me hell for about twenty minutes on the phone tonight, which bothered me no end. . . . I had been thinking of that this last time you rang the bell, and I had a very unpleasant feeling about it. . . . I was rehearsing their conversation in my mind was what was . . . going on.

The results for the three correspondence dimensions support the idea that the relationship of dream material to waking life is considerably less direct during REMPs than outside them. Incidentally, the one memory process noted in an REMP was the recreation of an event from early childhood, while other memory processes were clearly related to much more recent thoughts, events, or situations. Since the evaluation of reports on the first two of the correspondence dimensions involved a degree of judgment on the experimenter's part not found elsewhere in the analysis of the transcript data, it is important to note that the conclusions reached on the basis of this evaluation are substantiated by the subject's own ratings and comments reported below.

Dream Information Survey. Table 5 presents the data on those questions which successfully differentiated REMP and non-REMP reports. With a six-choice scale for the rating questions, scale values of 0 to 5 were arbitrarily assigned to the alternatives and mean scores computed.

As was the case with the transcript data it is clear that REMP dreams are characterized formally by greater organismic involvement (Emotion: Self, Emotion: Others, Anxiety, Proportion of scenes clearly visual, Activity, Violence-Hostility) and elaboration (Scenes, Dramatic, Clarity of content). The last-mentioned category had a 7–1–0 consistency figure, but its REMP versus non-REMP difference failed to attain statistical significance.

The correspondence data clearly support earlier conclusions that non-REMP dreams more often show a manifest relation to waking life, and that such relations are less distorted when they occur outside REMPs. The subjects' responses to the question "If [this dream] makes any sense to you, how does it make sense?" were coded to obtain the Prior experience-Recent proportion. Those replies mentioning any prior experience were classified as to whether this experience was recent (within a month of the dream day) or not. The month criterion was selected as it seemed a natural boundary, most reports relating to events well within that time or to ones considerably earlier in the subject's life. "No relation" was scored where the subject failed to indicate a *manifest* relation to some *specific* event, thought, or situation in his past history.

The Dream Information Survey results, therefore, seem to support the major conclusions drawn from the transcript data, both about formal and correspondence characteristics of sleep productions.

It should be noted here that while we have been dealing with statistically significant differences between REMP and non-REMP content, these differences in absolute terms are seldom gross. There is considerable overlap between results within and outside REMPs,

TABLE 5

SIGNIFICANT DIFFERENCES IN QUALITATIVE CHARACTERISTICS OF REPORTED PHENOMENA[a]

Dimension	REMP		Non-REMP		Consistency
	n	Mean score	*n*	Mean score	
Formal Differences					
Activity	87	2.10	68	1.38	7-1-0*
Emotion: Self	88	1.41	69	.91	6-0-2*
Emotion: Others	87	1.19	65	.72	7-0-1**
Anxiety	89	1.07	68	.62	6-1-1*
Scenes (Mean *n*)	87	(2.30)	68	(1.57)	7-1-0**
Proportion of scenes clearly visual	87	(.74)	68	(.52)	8-0-0***
Dramatic	87	1.10	70	.54	8-0-0***
Violence-Hostility	89	.71	70	.23	7-1-0**
Correspondence differences					
Distortion	87	1.68	69	.75	8-0-0***
How makes sense:					
Prior experience-Recent (prop.)	93	(.48)	95	(.73)	8-0-0***
No relation (prop.)	93	(.34)	95	(.22)	6-0-2*

[a] Source: Dream Information Survey.
* Significant at .05 level.
** Significant at .02 level.
*** Significant at .01 level.

and this overlap is as deserving of attention as are the differences.

DISCUSSION

Incidence of Recall Outside REMPs

Dement and Kleitman (1957b) have suggested that the few cases of dream recall from outside REMPs found in their study represented memories of preceding REMP dreams. These same authors (1957a) have suggested that the presence of spindles and/or delta waves is associated with the suspension of consciousness during sleep or at least with its maintenance at very low levels.

The Dement and Kleitman hypothesis that dream recall outside REMPs is delayed recall of prior REMP dreams seems improbable in light of the fact that in the present study the proportion of recall in Stages 2, 3, and 4 when there had been no intervening undisturbed REMP since the subject was last awake (.72, $n = 39$) was almost exactly comparable to the proportion of recall in these stages when there had been one or more such intervening undisturbed REMPs (.73, $n = 33$).

The considerable and consistent qualitative differences found between REMP recall and non-REMP recall in the present study also make the hypothesis that spindle and delta recall represents memories of prior REMP dreams seem improbable. It is difficult to imagine that the spindle and delta reports represent the end of a series of passive, memorial transformations of REMP material. In addition, when subjects were aware of reporting dreams which had occurred considerably prior to the bell (the seven cases mentioned previously), those dreams retained all the typical characteristics of REMP dreams, and this was true even when a considerable interval of time had elapsed since the last REMP. This last bit of evidence argues that REMP dreams persist, if they do so at all, relatively intact.

Two other ways in which spindle and delta dreams and thoughts could be something other than what they appear to be suggest themselves. Recall in such circumstances could be of hypnagogic phenomena experienced in low voltage, nonspindling EEG periods while falling asleep. On the other hand, it could represent hypnopompic phenomena (phenomena occurring during the state of awakening from sleep) experienced between arousal by the bell and full alerting.

The hypothesis that non-REMP recall is of hypnagogic material seems, on the basis of available evidence, to lack adequate empirical justification. The hypnagogic awakenings in the present study which produced reports similar to those from other non-Stage 1 awakenings were made *after* spindle and delta onset, and, although subjects least often felt for this category of awakenings that the phenomenon they reported was occurring just as the bell rang, they still felt on 83% of these occasions that such contemporaneity had obtained. Furthermore, despite broad similarities, there were enough differences between the hypnagogic reports and reports from later spindle and delta awakenings to suggest that those later reports were not unretouched, or slightly faded, memories of hypnagogic material. A theory that stated that pre-REMP Stage 2 recall consisted of memories of hypnagogic material might have some difficulty, for example, in explaining the finding that manifest work-school themes were present on 50% of 26 Stage 2 awakenings which produced recall but on only 23% of 26 recall producing hypnagogic awakenings. Finally, two of the pilot subjects were run for several evenings with a number of hypnagogic awakenings made on each such occasion. These awakenings were made at points ranging from the break-up of continuous alpha to several minutes after spindle onset. Awakenings made during low voltage EEG activity on falling asleep seemed to produce quite different material (e.g., "floating" sensations, "flashing lights," "lantern slide" phenomena, fleeting progressions of thoughts and images) than that observed after spindle onset (more extended and self-involved "dreamlets"). This strengthens the conclusion that neither the hypnagogic reports nor the whole class of spindle and delta reports collected in this study consist of memories of material which occurred in Stage 1 while falling asleep.

It is more difficult to rule out the hypothesis that non-REMP recall represents hypnopompic imagery, since we really know very

little about the subjective aspects of waking up. Nevertheless, there are some lines of evidence to recommend that we also reject this hypothesis. On those awakenings for which arousal took the longest time, recall was very rarely obtained. Unless immediate contact could be made with an alerted subject, the chance of obtaining some substantive material was slim. Thus if our spindle and delta recall was of hypnopompic material, that material had to transpire in a very brief moment of time. But, although median estimated dream lengths were usually shorter for non-REMP categories of awakening than for REMP categories, they were not always so, and in no case did they vary nearly so dramatically from the REMP figures as did the time of non-REMP arousal from the average length of the REMP. Complete arousal after Stage 2 awakenings took only a matter of a few seconds, yet dreams from these awakenings were generally estimated to be as long as, and were described as multipart dreams almost as often as, dreams from awakenings made after REMPs of 9 or more minutes duration. The hypothesis under consideration might also have some difficulty in explaining why the material obtained from one category outside REMPs differed along certain dimensions from that from another such category. While we cannot definitely reject it, the weight of available evidence seems to run counter to the hypnopompic hypothesis. It appears, then, that the non-REMP recall obtained in the present study is not artifactual in nature.

A Reconciliation with Previous Studies

If the spindle and delta dreams and thoughts reported in this study represent phenomena which actually occurred contemporaneously with these non-REMP brain wave patterns, as now seems likely, we are faced with an apparently very difficult situation—how can we reconcile the present results with those of recent studies by investigators of the Chicago group which unanimously seem to indicate that dreaming is very unlikely to occur, if at all, outside Stage 1 REMPs? It appears that several factors contributed to the high incidence of recall of mental phenomena at awakenings outside Stage 1 reported in this study and at variance with the results of recent Chicago studies.

Criterion of Recall. Dement and Kleitman (1957b) state that "vague, fragmentary impressions of content were considered negative" (p. 341). Such fragmentary impressions were here counted as content, and since such material was much more likely outside REMPs, it partially explains the high level of non-REMP recall in this study. Some pertinent evidence is available on the amount of the discrepancy between the results of this study and those of other Chicago studies which can be accounted for by differing criteria of recall. Orlinsky, in an unpublished analysis of the extensive records collected by Kamiya, found that the application there of the criterion used in this study produced results indicating recall in over 40% of the awakenings made outside Stage 1 REMPs. It was only when a standard of greater detailedness of content was employed that non-REMP recall incidence dropped to a level comparable to that in the Dement and Kleitman work.

Interest in detailed, coherent dreams seems quite understandable, but the labeling of "vague, fragmentary impressions of content" as "no recall" is somewhat unfortunate if the recall results are then used to argue that spindles and delta are very likely associated with a suspension of conscious mental activity. It now seems time to embark upon systematic studies of *all* aspects of sleep consciousness, whether they are "dreamlike" or not.

Subject Set. Some protocols suggest that asking whether anything was "going through your mind" elicited material from subjects which might have been missed had they only been asked whether they were "dreaming." Subjects demonstrated that they considered dreams to be visual, and to be somewhat bizarre. Since non-REMP reports were less often visual and had a higher degree of correspondence with reality, the application of these criteria by subjects would serve to depress recall outside REMPs most severely if the subject believed that the experimenter was interested in his "dreams," and in these phenomena alone.

Continued Subject Arousal. It appears that there were at least 10 non-REMP instances of recall in the present study and 1 REMP instance which might have been missed had not the continued arousal technique been employed. The clear and direct way in which the material seemed to come into consciousness in these cases, and the subject's reactions to the material, strongly suggested that memory, and not fabrication, processes were responsible for the production of the material in question.

Nature of Content. The comments of one subject suggested that not only is material which has a high degree of correspondence less likely to be labeled as a "dream," it is perhaps less likely to be recalled at all. If it is true that unusual material is likely to be better recalled, this would clearly favor REMP recall and depress non-REMP recall.

One additional factor which makes the disparity between this study and other recent studies appear less formidable is that the present study has produced results in accord with the conclusions of a number of largely unsystematic but important studies conducted a number of years ago (Blake, Gerard, & Kleitman, 1939; Davis, Davis, Loomis, Harvey, & Hobart, 1938; Henry, 1941; Knott, Henry, & Hadley, 1939). The results of the present study are a striking confirmation of the results of the most systematic of these studies, that of Teplitz (1943). Teplitz found that reports of having dreamed were obtained in 63% of 78 spindle and delta awakenings, and concludes "that dreams may occur in association with any type of sleep potential pattern" (p. 12). It is hoped that the present study has pointed up the need for a reconciliation of earlier and of more recent electroencephalographic studies of dreaming. In this study's suggestions as to why the incidence of non-REMP "dreaming" has been understated in recent studies, it is also hoped that the way has been pointed toward such a reconciliation.

Present Results and Freudian Theory

If we may view the data from awakenings around REM onset as pertinent to "dream formation" or at least to the "formation" of that class of dreams with which Freud was most concerned, they tend to provide striking confirmation of certain aspects of Freud's theory of dream construction. Freud (1956) apparently did not feel that stages in dream development would be represented in the dreamer's consciousness (p. 576). The present data, from material which proved to be "available" to the dreamer, might be thought of as providing some measure of support for the idea that day residues are "worked over" as dream development, or rather REMP dream development, progresses. Although processes of distortion are active in Stage 2, and processes preserving correspondence are active in Stage 1, there are shifts in the incidence of characteristics of dream content between these two stages which mirror Freudian conceptions of dream development.

SUMMARY

Eight subjects were run for a total of 57 nights: to test the hypothesis that dreams, or some form of conscious mental activity, occur in all stages of sleep and ocular activity; if this hypothesis proved correct, to explore possible systematic differences in form and content of reports obtained from awakenings in different stages of sleep and ocular activity; and in any event, to investigate possible qualitative progressions and their EEG and eye movement correlates in the "formation" of dreams occurring in EEG Stage 1 rapid eye movement periods (REMPs).

The mean proportion of recall for all subjects was .74 in non-REMP awakenings and .87 in REMP awakenings. Pooled subject data showed that recall occurred in 70% of the awakenings made during deep sleep, a period of the sleep cycle previously felt to be quite incompatible with dreaming. It is suggested on the basis of the present results that reportable mental activity is always present in the sleeping human.

Reports obtained in periods of REM activity showed more organismic involvement in affective, visual, and muscular dimensions and were more highly elaborated than non-REMP reports. REMP reports showed less correspondence to the waking life of the subject than did reports from spindle and delta sleep. The relatively frequent occurrence of thinking

and memory processes in spindle and delta sleep was an especially striking result.

REFERENCES

BLAKE, HELEN, GERARD, R. W., & KLEITMAN, N. Factors influencing brain potentials during sleep. *J. Neurophysiol.*, 1939, **2**, 48–60.

DAVIS, H., DAVIS, P. A., LOOMIS, A. L., HARVEY, E. N., & HOBART, G. Human brain potentials during the onset of sleep. *J. Neurophysiol.*, 1938, **1**, 24–38.

DEMENT, W. Dream recall and eye movements during sleep in schizophrenics and normals. *J. nerv. ment. Dis.*, 1955, **122**, 263–269.

DEMENT, W., & KLEITMAN, N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *EEG clin. Neurophysiol.*, 1957, **9**, 673–690. (a)

DEMENT, W., & KLEITMAN, N. The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *J. exp. Psychol.*, 1957, **53**, 339–346. (b)

DEMENT, W., & WOLPERT, E. A. Relationships in the manifest content of dreams occurring on the same night. *J. nerv. ment. Dis.*, 1958, **126**, 568–578. (a)

DEMENT, W., & WOLPERT, E. A. The relation of eye movements, body motility, and external stimuli to dream content. *J. exp. Psychol.*, 1958, **55**, 543–553. (b)

FOULKES, W. D. Dream reports from different stages of sleep. Unpublished doctoral dissertation, University of Chicago, 1960.

FREUD, S. *The interpretation of dreams.* New York: Basic Books, 1956.

GOODENOUGH, D. R., SHAPIRO, A., HOLDEN, M., & STEINSCHRIBER, L. A comparison of "dreamers" and "nondreamers": Eye movements, electroencephalograms, and the recall of dreams. *J. abnorm. soc. Psychol.*, 1959, **59**, 295–302.

HENRY, C. E. Electroencephalographic individual differences and their consistency: I. During sleep. *J. exp. Psychol.*, 1941, **29**, 117–132.

KNOTT, J. R., HENRY, C. E., & HADLEY, J. M. Brain potentials during sleep: A comparative study of the dominant and non-dominant alpha groups. *J. exp. Psychol.*, 1939, **24**, 157–168.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

TEPLITZ, ZELDA. An electroencephalographic study of sleep and dreams. Unpublished master's thesis, University of Illinois, 1943.

WOLPERT, E. A. Studies in psychophysiology of dreams: II. An electromyographic study of dreaming. *Arch. gen. Psychiat.*, 1960, **2**, 231–241.

WOLPERT, E. A., & TROSMAN, H. Studies in psychophysiology of dreams: I. The evocation of sequential dream episodes. *Arch. Neurol. Psychiat.*, 1958, **79**, 603–606.

(Received April 21, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 26–31

THE COMPETITION OF AFFECTIVE RESPONSES IN HUMAN SUBJECTS

GERALD A. MENDELSOHN [1]

University of California, Berkeley

Competing response or interference hypotheses have been used to explain a wide range of behaviors. For example, Guthrie (1935) views extinction as the result of the acquisition of a new response incompatible with the old response to a given stimulus. Likewise, Osgood (1953) stresses the importance of interference effects in serial learning and Miller (1944) analyzes the effects of competing responses in conflict behavior. In another experiment, Ullman (1951) shocked rats in the same compartment in which they had previously been given food. After an initial reduction, eating increased especially during the period when the shock was on. The author interprets this result as evidence that eating becomes a generalized tension reducer in this situation. He argues that this position is supported by the fact that even satiated animals ate when the shock was applied. A process of affective response competition can be inferred from this result, i.e., the "tension reducing" character of eating competed with the shock arousal, attenuating its painfulness.

Certain clinical observations also suggest that a competing response hypothesis can be extended to the area of affective processes. Fenichel (1939) describes the paradoxical behavior of counterphobic individuals, patients who, rather than avoiding anxiety inducing stimuli, seek them out. However, while exposing themselves to situations of which they are or were afraid, such individuals do not experience feelings of anxiety. Fenichel suggests that the absence of anxiety in this circumstance is sometimes due to the simultaneous appearance of reassuring aspects in the situation. Thus, the patient may attempt to perform an otherwise frightening act in the company of a comforting or protecting person. Once again a process of affective competition can be inferred from the resulting attenuation or elimination of anxiety.

The present experiment is an attempt to test the applicability of a competing response hypothesis to human affective processes. The specific hypothesis is as follows: a reduction of the magnitude of a conditioned anxiety response can be produced by the simultaneous elicitation of certain other affects. The affects chosen were contentment, delight, and sadness. The first two were included in order to study the interaction of two different positive affects with the negative affect of anxiety and sadness was chosen in order to study the interaction of a different negative affect with anxiety.

The major difficulty facing the experimenter in research of this type is that of finding some method of producing the conditions necessary to test the competing affects hypothesis. Conditioning an unpleasant affective response to a stimulus is a relatively easy task, but it is much more difficult to condition pleasant affective states to particular stimuli. Since the individuals used in the research were all well-trained hypnotic subjects, it was possible, by suggestion during the trance state, to attach a variety of responses to the experimental stimuli. The hypnotic induction was intended to serve only one purpose—to form the association between the stimulus and the affective response. To avoid complications arising from the nature of the hypnotic state itself, all parts of the experiment, except the inductions, were conducted with the subject fully awake.

The hypnotic training and the inductions were done by staff members of the Department of Psychology at the University of Michigan.

[1] This study was done in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan in 1959. The writer wishes to express his thanks to Gerald S. Blum for his guidance in designing and carrying out the experiment. The research project of which the present study is one aspect, was supported by Ford Foundation and United States Public Health Service grants.

PROCEDURE

Apparatus

A Gerbrands-Dodge type tachistoscope, designed for use with individual subjects, was employed. The stimuli were presented at a speed sufficiently slow (.15 second) and at illuminations sufficiently high to allow for perfectly accurate perception. The illumination of the adapting field was 1.205 foot Lamberts and of the viewing field, 4.135 foot Lamberts. The speed and illuminations were arrived at on the basis of pretest data which indicated that subjects could correctly identify the stimuli at the levels reported above.

Subjects' galvanic skin response (GSR) was measured by a Fels dermohmeter. One zinc electrode was placed on the subject's left palm and the other at the midpoint between the elbow and the wrist of the left arm.

Dependent Variable

Anxiety was chosen as the negative affect in this experiment because it is easily measured by objective means. Increases in conductance in the subject's GSR were used as the dependent variable, the measure of the presence and intensity of anxiety. Previous unpublished experimentation in connection with the project of which the present study is one aspect, indicated that only the affect of anxiety was regularly associated with GSR activity. Consequently, where other affects were elicited as part of the procedure, there was little chance that they would directly augment the amplitude of the GSR deflections (see Table 3 and accompanying discussion).

The GSR data are in the form of conductance in millimhos (mmhos). A deflection was scored only if it began not more than 4 seconds after the presentation of the stimuli.

Subjects

Subjects were six male students of 21 or more years of age. All were paid volunteers who had participated in several experiments.[2] Before any experiments were conducted, the subjects spent several hours in hypnotic training. During this period, the suggestibility of each subject was evaluated. Of 13 original candidates, 10 were carried through to completion. The other 3, while hypnotizable, did not enter into sufficiently deep trances and were discontinued prior to any experimentation. The present study was initiated after 4 experimental subjects had been terminated. It is important to note that at the time they volunteered, the subjects did not know hypnosis was to be used as part of the experimental procedure. Because of the selection methods used, there is reason to believe that the present subject group is not radically different from a typical group of volunteer subjects for a psychological experiment.

[2] The other experiments were of considerable variety. While some involved the induction of anxiety, none involved any of the other affects used in the present experiment, nor was the competition of affects part of any previous study.

Design

Five letters, ¾ inch high and ⅝ inch wide, printed with India ink on 9 inch × 10.5 inch stimulus boards, were used as stimuli. The letters were H, S, C, D, and N. A different affective response was conditioned to each of these letters, except the letter N, by the following hypnotic induction:

> Whenever you see the letter C, whether awake or asleep, you will feel very content. Whenever you see this letter you will have a feeling of contentment.

This induction was repeated for all the letters, each in turn being connected with its specific affect. The order in which this was done is as follows: C and contentment, S and sadness, D and delight, and H and anxiety. (H was selected as the conditioned stimulus for anxiety because it produced no common abbreviations or initials when paired with the other letters.) No mention was made of the letter N during hypnosis. If difficulty was encountered in conditioning the affective response to the stimuli, the induction was repeated and strengthened. In order to test the efficacy of the induction while the subject was still under hypnosis, he was asked to report his subjective feeling when shown the letters individually and in the case of the letter H, his GSR was noted. After the initial induction, supplementary inductions were made when necessary. These were identical to the initial induction.

After the subject was awakened, the letters were shown two at a time on the stimulus boards, separated by 3 inches. Between trials the subject viewed a blank, white adaptational field with a fixation point in its center. During the trial, the two letters appeared on a line with and 1.5 inches to the left and right of the fixation point. H appeared on every trial, half the time as the letter on the right and half the time as the letter on the left. Every possible combination of H with the other letters was shown an equal number of times. The stimuli were shown in a tachistoscope for .15 second, a duration which allowed for perfect identification of the stimuli. The subject was given the following instructions:

> I am going to show you two letters in the tachistoscope, one on the left and one on the right. But before giving you further instructions, I will show you the stimuli you are going to see.

The subject was then shown the five letters, one at a time, and if the letter H led to a GSR deflection, the instructions were continued:

> When I say "Ready," I want you to focus on the left. It is important to remember this—focus on the left on every trial. I will remind you of this every four trials. After the flash, I will ask you first what you saw on the right and then I will ask you what was on the left.

The subject was told to focus on the left for all presentations so that the experimenter could be reasonably sure of which stimulus the subject saw first. Thus, it was likely that on half the trials, the anxiety stimulus was perceived first and on the other half, one of the other affective stimuli was perceived first.

The check series consisted of presentations of each letter individually. The letters in this series were the same size as those in the experimental series, but they were printed in the center of the stimulus board. This series was run to check on the efficacy of the induction. It was expected that H would always lead to a GSR deflection and that the other stimuli would not consistently do so. There were four check series, one preceding each of four blocks of experimental trials.

The test series was run in eight blocks of eight trials, each stimulus combination appearing once in each block. The schedule of presentations was arranged in such a manner that each pairing appeared equally often in the first, second, third, and fourth quarters of the blocks of eight. Within these limits, the order of presentation was established by reference to a table of random numbers.

After the experiment was finished, an inquiry was conducted in which the subject was asked to rank the stimulus combinations along a subjective unpleasantness dimension.

In line with the competing affects hypothesis, it was predicted that the pairings of anxiety and contentment, sadness and delight would produce a level of GSR deflection that differed from that produced by the pairing of the anxiety with the neutral stimulus. It was reasoned that contentment would have an antagonistic effect on anxiety since these two affects are quite dissimilar, but that delight and sadness are similar to anxiety in that all three reflect states of arousal. Thus, in the case of the pairing of anxiety and contentment, a lower magnitude of GSR deflection was predicted, but no directional prediction was made for the other two combinations.

RESULTS

The total change in conductance for each stimulus combination for each subject is presented in Table 1. In every case the combination of the contentment (C) and anxiety (H) stimulus produces less GSR activity than the control condition, the pairing of anxiety and the neutral (N) stimuli. There is no consistent pattern for the other two pairings.

A Rows × Columns × Blocks analysis of variance was run on the GSR data for the subjects as a group. The fixed effects used were stimulus combination and order (the appearance of the anxiety stimulus on the left or the right). The combination main effect was evaluated by the Combination × Subject interaction, the Order Effect × the Order × Subject interaction. The results are summarized in Table 2.

TABLE 1

TOTAL INCREASE IN CONDUCTANCE IN MILLIMHOS FOR EACH STIMULUS COMBINATION FOR EACH SUBJECT

Subject	Condition			
	N	S	C	D
A	3.93	2.36	.50	3.32
B	8.32	9.08	3.76	6.32
C	.75	.64	.62	.69
D	3.72	4.08	1.91	4.99
E	.46	.13	.14	.05
F	1.74	4.46	1.62	2.07
Sum	18.92	20.75	8.55	17.44

The *F* ratio for the combination main effect is significant at the .05 level but inspection of the means for the four combinations indicates that the result is attributable primarily to the HC pairing. The *t* test between this combination and any other combination is significant at the .01 level or beyond in each case. (The Combinations × Subjects interaction mean square was used to derive the error term for the *t* tests.) No other significant differences between the means were found. This analysis strengthens the impression that only the combination of H and C leads to a decrease in the magnitude of GSR deflections. The insignificant Conditions × Subjects interactions indicates that this effect is consistent across subjects.

The argument can be raised that the above results can be explained by GSR interaction rather than by a central, affective interaction. That is, perhaps the neutral, sadness, and de-

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE FOR STIMULUS COMBINATIONS

Source	*df*	*MS*	*F*
Combination (C)	3	2.4438	3.7545*
Order (O)	1	2.7840	4.5707
Subjects (S)	5	11.5156	
C × O	3	.8207	<1
C × S	15	.6509	<1
O × S	5	.6091	<1
C × O × S	15	1.2066	

* Significant at .05 level.

light stimuli alone produce more GSR activity than the contentment stimulus alone and so when they are combined with the anxiety stimulus, they produce a greater mean GSR deflection. An examination of the check trials makes this argument untenable. The contentment stimulus shown individually produced a mean deflection of .56 mmhos per subject; the neutral stimulus, .63 mmhos; the sadness stimulus, .46 mmhos; and the delight stimulus, .56 mmhos. The anxiety stimulus produced a mean deflection of 1.30 mmhos per subject. An analysis of variance run on the check trial data is summarized in Table 3.

The significant stimulus main effect is due to the anxiety stimulus alone, for subsequent t tests indicate that it produces significantly greater increases in conductance ($p = .025$ or beyond) than any other stimulus while the GSR deflections associated with the other stimuli are not significantly different from each other. A significant difference at the .05 level requires a mean difference of .55 mmhos.

In view of these results, the competing affects hypothesis is confirmed for the pairing of contentment and anxiety, in that the level of GSR conductance increases associated with this combination is significantly lower than that associated with the pairing of the neutral and anxiety stimuli. Further, this appears to be due to a process of affective interaction rather than some sort of direct GSR summation. In passing it can be pointed out that GSR deflections, while associated with affective states other than anxiety, are greater in magnitude and more consistently evoked in the presence of the anxiety stimulus.

The data for the subjective report of the subjects is presented in Table 4. Kendall's coefficient of concordance for the data indicates that there is high intersubject agreement

TABLE 3

SUMMARY FOR ANALYSIS OF VARIANCE OF STIMULI SHOWN INDIVIDUALLY

Source	*df*	*MS*	*F*
Stimuli	4	.6934	3.8501*
Subjects	5	1.2971	7.2021**
Residual	20	.1801	

* $p < .05$.
** $p < .01$.

TABLE 4

RANKINGS OF THE SUBJECTIVE UNPLEASANTNESS OF THE STIMULUS PAIRINGS BY EACH SUBJECT

Subject	Condition			
	N	S	C	D
A	2	1	3	4
B	3	2	4	1
C	1	3	4	2
D	2	1	3	4
E	2	1	3	4
F	2	1	3	4
Sum	12	9	20	19

on the ranking of the four combinations ($p < .05$). This is in line with the absence of a significant Subject × Combination interaction reported above. It is interesting to note that the order of the subjective rankings, S, N, D, C, is the same as the order produced by the GSR data.

DISCUSSION

The results of the competing affects experiment indicate that a reduction of the magnitude of conditioned anxiety responses can be produced by the simultaneous elicitation of certain other affects. However, of the affects used, only contentment led to a consistent reduction of the anxiety response. This finding is somewhat difficult to explain, for the interaction of the two positive affects, contentment and delight, with anxiety produced different GSR results. In the light of this finding, the simple formulation that emotions of opposite tone, negative and positive, have an antagonistic effect is untenable. Consequently, it is necessary to consider the difference between contentment and delight and in so doing to develop some hypotheses about the nature of affect.

Jacobson (1953) suggests that there are two kinds of pleasurable affects, those which arise from increases in "tension" and those which arise from decreases in tension (relief pleasures). Assuming that this is the case, the processes underlying the two forms of pleasurable affect are quite different. The first class, tension pleasures, includes those affects which accompany increases in arousal and excitation while the second class includes those states which accompany the reduction

of arousal or excitation. The intensity of the affect is in both cases a reflection of the magnitude of change in excitation. It appears that in the experiment under consideration, the feeling of delight can be described as a tension pleasure, while contentment can be described as a relief pleasure. Thus, the difference between the two is one which corresponds to a basic difference in the processes underlying the production of each. Further, if the argument is extended to negative, unpleasant affects, it can be hypothesized that they correspond to increases in arousal which exceed some upper threshold.

If the results of the experiment are considered in the light of this reasoning, a consistent explanation can be formulated. Anxiety and sadness can be described as arousal states in which the level of excitation exceeds the hypothesized upper threshold, while delight can be described as an arousal state in which the level of excitation is below the threshold. Only contentment would correspond to a decrease in excitation and finally the neutral stimulus would have no tension affecting qualities at all. Thus, when anxiety and sadness or delight are combined there is an interaction of arousal states leading to a high level of excitation and, thus, of GSR. However the simultaneous arousal of anxiety and contentment may lead to antagonistic responses, one tending to increase and the other to decrease excitation. The resultant of these processes is a decrement in arousal and, thus, less GSR and subjective unpleasantness.

To summarize, the order of magnitude of GSR drops for the four stimulus combinations is explained by the hypothesized underlying arousal state dynamics. When sadness and anxiety are combined there is an interaction of two above threshold arousals, thus, the highest level of GSR and subjective unpleasantness. When the neutral condition and anxiety are combined, there is no interaction, the second highest level of GSR and subjective unpleasantness being attributable to the above threshold arousal of anxiety alone. When the delight and anxiety states are combined there is once again an interaction of arousal states but one is above threshold and the other below threshold. The resultant is a level of GSR below that of the control condition (HN) but well above that of the only combination of affective stimuli, anxiety and contentment, in which there are antagonistic arousal states.

In developing this interpretation certain hypotheses about the nature of affect have been made: first, that all affects are the result of the same quantitative factor, excitation or arousal; second, that the pleasure and pain qualities of affects are determined by the level of excitation; and third, that affect is basically a central phenomenon, peripheral reactions like the GSR being responses emitted in consequence of the central state. The theory underlying this interpretation, then, is an arousal theory similar to those of Arnold (1950), Lindsley (1951), and Jacobson (1953).

SUMMARY

Competing response hypotheses have been employed to explain a wide range of behaviors. The present study is an attempt to study competing responses in human affective behavior. The following hypothesis was tested: a reduction of the magnitude of a conditioned anxiety response can be produced by the simultaneous elicitation of certain other affects.

The procedure for the experiment consisted of the simultaneous tachistoscopic presentation of two letters, one having an anxiety response conditioned to it under hypnosis, and the other a response of either sadness, contentment, or delight. In the control condition the other letter was affectively neutral.

The competing affects hypothesis was confirmed for the pairing of contentment and anxiety in that the magnitude of GSR deflection associated with this combination was significantly lower than that associated with the pairing of the neutral and the anxiety stimulus. The pairing of anxiety and sadness produced greater GSR drops than any other condition, while the pairing of anxiety and delight produced slightly less GSR activity than the control condition, though significantly more than the pairing of anxiety and contentment. Evidence was presented to suggest that these results are due to central affective interaction rather than peripheral GSR interaction.

REFERENCES

ARNOLD, MAGDA. An excitatory theory of emotion. In M. Reymert (Ed.), *Feelings and emotions*. New York: McGraw-Hill, 1950. Pp. 11–33.

FENICHEL, O. The counterphobic attitude. *Int. J. Psycho-Anal.*, 1939, **20**, 263–274.

GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.

JACOBSON, EDITH. The affects and their pleasure-unpleasure qualities, in relation to the psychic discharge processes. In R. Loewenstein (Ed.), *Drives, affects and behavior*. New York: International Univer. Press, 1953. Pp. 38–66.

LINDSLEY, D. Emotion. In S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 473–516.

MILLER, N. E. Experimental studies in conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. 1. New York: Ronald, 1944. Ch. 14.

OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.

ULLMAN, A. D. The experimental production and analysis of a "compulsive eating symptom" in rats. *J. comp. physiol. Psychol.*, 1951, **44**, 575–581.

(Received April 22, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 32–40

THE EFFECTS OF COMPETITION ON REACTION TIME AND PALMAR SKIN CONDUCTANCE[1]

RUSSELL M. CHURCH

Brown University

The early work on social facilitation (Allport, 1924; Dashiell, 1930; Triplett, 1897) suggested that social factors such as the presence of co-workers, rivalry, and competition may increase the speed of response. The present study began as a further investigation of the effect of a competitive situation on the speed of response. The response chosen was the simple reaction time, a standardized laboratory task in which latency can be accurately measured.

A competitive situation may be defined as one in which an individual's success is determined by some characteristic of his response relative to that of another individual. In the present experiment subjects were always tested in pairs. The instructions for both the normal and competitive condition emphasized the speed of response, but the instructions for the competitive condition also emphasized the speed of response relative to that of the other subject. On each trial under the competitive condition, unlike the normal condition, a stimulus was presented as soon as the first subject had responded indicating which subject had responded with the shorter latency. The first problem was to determine whether such a competitive situation would result in a decrease in the latency of the simple reaction time.

There are a number of possible explanations of competitive facilitation. If the competitive situation results in a decrease in response latency, an attempt should be made to determine which aspects of the competitive situation are responsible for the facilitation. If the competitive facilitation is gradual, a plausible hypothesis is that the competitive situation results in differential reinforcement of fast responses, and that this differential reinforcement of fast responses results in decreased reaction times. It is clear that in any competitive situation there will be differential reinforcement of some response characteristic, unless one subject receives all of the reinforcements. If the faster of two subjects were reinforced on each trial, then each subject should be reinforced more often for his fast responses than for his slow responses. Differential reinforcement of fast responses has resulted in faster rates of response of animals in the Skinner box (Ferster & Skinner, 1957), and Church (1961) has demonstrated that differential reinforcement was responsible for facilitation of response rate of pairs of rats competing for food in lever boxes. The differential reinforcement of fast responses that occurs in a competitive situation may result in faster reaction times of human subjects.

If the competitive facilitation is abrupt, and particularly if it appears on the first trial, the differential reinforcement hypothesis is implausible. One explanation of an abrupt facilitation of the reaction time in a competitive situation is based on previous learning, e.g., subjects have previously learned to respond quickly in competitive situations involving speed of response and thus respond quickly after the competitive cues are introduced. An alternative explanation is motivational, e.g., competition, for some reason, increases the level of motivation which, in turn, decreases the latency of the simple reaction time. The second problem of the first experiment, then, was to determine whether the decrease in the latency of the simple reaction time in a competitive situation was gradual or abrupt.

Experiment I: Competitive Facilitation of the Simple Reaction Time

Method

Subjects. Subjects were 36 college students without previous experience with a reaction time apparatus. They were volunteers from small sections of an introductory psychology course and were paid $1.00

[1] This investigation was supported by a PHS Research Grant, M-2903, from the National Institute of Mental Health, United States Public Health Service.

for participation in a single 50-minute session. There were nine pairs of men and nine pairs of women.

Apparatus. The apparatus consisted of two reaction time devices, mounted on desks side-by-side. Each device had two 7-watt lights (green and amber) and two toggle switches (right and left), although in the present experiment only the green light and the right switch were used. The toggle switches were 8 inches apart, and the lights were mounted in the vertical plane halfway between the two toggle switches. A 6 volt ac, 120 cycle tone through a loudspeaker served as the ready signal. Anticipatory responses, e.g., responses during the ready signal, were indicated by a bell for one subject and by a buzzer for the other subject. The duration of the ready signal, the onset of the light, and the intertrial interval were timed by automatic equipment. The experimenter could control whether the light of each subject would go off immediately following his response (the normal condition) or whether the light of the first subject to respond would go off immediately and the light of the other subject would remain on for 6 seconds providing information to both subjects regarding the winner of the trial (the competitive condition). Reaction times were recorded in hundredths of a second (Standard Electric timers, Model A-1).

Procedure. Prior to the experiment subjects were simply told that the experiment involved the reaction time, and that they would come to the experimental room in pairs. At the time of the experiment, subjects were instructed

> This is an experiment on the reaction time. Your task is to pull this toggle switch toward you as fast as you can after the light goes on. On each trial there will be a tone that is the ready signal. It will go on for somewhere between 1 and 10 seconds, then the light will go on. Your task is to pull the toggle switch toward you as fast as you can as soon as you see the light. On each trial, I will tell you when to reset the switch. Do you have any questions?

1. Trials 1–50. Fifty trials were then given. The length of the ready signal was a uniform distribution from 1 to 10 seconds randomly selected for each pair of subjects. The light came on simultaneously with the termination of the ready signal, and it went off when the subject moved the toggle switch. (It should be noted that there were clearly audible stimuli simultaneous with the light onset, so the stimulus for the reaction had auditory as well as visual properties.) The time between light onset and the beginning of the next ready signal was 10 seconds. After 50 trials the experimenter said, "Let's take a break."

2. Trials 51–100. After a 2-minute rest period the second block of 50 trials was begun. For half the subjects the situation was the same as during the first block of trials, and the experimenter said

> You have done one block of 50 trials. Let's do a second block of 50 trials the same way.

The other half of the subjects were given competitive instructions. The experimenter said

> You have done one block of 50 trials. For this second block of 50 trials, I would like to make the following change. This time, instead of just going as fast as you can, try to go faster than your partner. If you do go faster, your light will go out. If you are slower, your light will stay on for 6 seconds. Do you have any questions?

The second block of 50 trials was then given. Again the length of the warning signal was a uniform distribution from 1 to 10 seconds, randomly selected for each pair of subjects. The time between light onset and the beginning of the next ready signal was 10 seconds. The treatment of the subjects in the normal condition, then, was the same as during the first block of trials. The treatment of the subjects in the competitive situation was the same, except that the light of the subject making the faster response went out immediately while the light of the slower subject remained on for 6 seconds. After 50 trials the experimenter said, "Let's take another break."

3. Trials 101–150. After a 2-minute rest period, the third block of 50 trials was begun. For the subjects who had the normal condition in the second block and the competitive condition in the third block, the experimenter said

> You have done two blocks of 50 trials. For the third block of 50 trials, I would like to make the following change. This time, instead of just going as fast as you can, try to go faster than your partner. If you do go faster, your light will go out. If you are slower, your light will stay on for 6 seconds. Do you have any questions?

For the subjects who had the competitive condition in the second block and the normal condition in the third block, the experimenter said

> You have done two blocks of 50 trials. For the third block of 50 trials, I would like to return to the first condition. Your task is to pull the toggle switch as fast as you can as soon as you see the light. The light will go out every time, but respond as quickly as you can.

Design. To summarize the procedure, each subject had 150 trials of the simple reaction time, divided into three blocks of 50 trials. Two of the blocks were under the normal (N) (noncompetitive) conditions and one of the blocks was under a competitive (C) condition. Nine pairs of subjects randomly selected, were given the order NCN; the other nine pairs were given the order NNC.

Results

The reaction time (RT) on each trial in hundredths of a second was transformed to log (RT − .10), .10 being considered an approximation to the irreducible minimum for this situation (Woodworth & Schlosberg,

1954, p. 21). All statistical tests were performed on this transformation, which served to normalize the distribution of reaction times and to stabilize its variance. Values of the "mean reaction time" in the remainder of this paper will refer to the antilog of the mean transformed reaction time, plus .10 second.

Figure 1 shows the mean reaction time as a function of blocks of 10 trials. Inspection of this figure suggests that the effect of competition was immediate rather than gradual. In both groups there was an immediate decrease in latency when the competitive condition was begun. The difference between the competitive and normal conditions, however, did not increase with trials. Figure 1 also suggests that the competitive condition temporarily decreased the reaction time during the subsequent normal condition.

Individual differences in the mean reaction time were large and reliable. The correlation between the mean reaction time during the first block of trials under the normal condition and the second block of trials under the normal condition was $r = .84$, for the NCN group and $r = .86$ for the NNC group. The correlation between the mean reaction time during the first block of trials under the normal condition and the first block of trials under the competitive condition was $r = .81$, for the NCN group and $r = .81$ for the NNC group. To reduce this source of variance, the mean reaction time of each subject during the first 50 trials was subtracted from his mean reaction time during the second and third blocks of 50 trials. An analysis of variance on these data, an analysis of two measurements on each of two independent groups, demonstrated that the reaction time was lower in the competitive condition than in the normal condition ($F_{1,34} = 62.7$, $p < .01$). Differences between the order of the treatments were slightly above the 5% level of confidence.

The Kolmogorov-Smirnov statistic was computed for each subject to determine if his reaction times in the competitive condition were stochastically smaller than those in the second normal condition. Sixteen of the 36 subjects, e.g., almost 50% of the subjects, were significantly faster (at the .01 level of confidence) in the competitive situation than in the normal situation. No subject was significantly faster (even at the .05 level of confidence) in the normal situation than in the competitive situation.

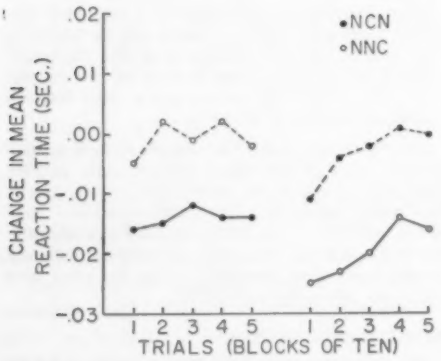

FIG. 1. Change in the mean simple reaction time as a function of blocks of 10 trials under the normal and competitive conditions—Experiment I. (Solid lines indicate the competitive condition; broken lines indicate the normal condition.)

EXPERIMENT II: COMPETITIVE FACILITATION OF THE DISCRIMINATION REACTION TIME

The first experiment demonstrated that a competitive situation may decrease response latency of a simple reaction. The second experiment was undertaken to determine whether a competitive situation also would decrease the response latency of a discrimination reaction.

Method

Subjects. Subjects were 18 of the students used in the previous experiment. For each subject the present experiment followed the former by exactly one week.

Apparatus. The apparatus was the same one used in the previous experiment, except that both lights and both switches were used and errors (responses with the wrong toggle switch) were also recorded.

Procedure. The instructions were the same as those used in the previous experiment, except for the sentence describing the task which was modified to read

> Your task is to pull this [the right] toggle switch toward you as fast as you can when the green light goes on, and to pull this [the left] toggle switch toward you as fast as you can when the amber light goes on.

The procedure was the same as that used in the previous experiment except that the stimulus light

used on each trial was determined randomly with each light having an equal chance of occurrence on any trial. Errors were indicated in the same manner as anticipatory responses, e.g., a bell or buzzer sounded, depending on the subject who made the error.

Design. The design was the same as that used in the previous experiment with the simple reaction time. Again two blocks were under the normal (N) (noncompetitive) conditions and one of the blocks was under the competitive (C) condition. Five pairs of subjects were given the order NCN; the other four pairs were given the order NNC. Half of the subjects had the same order of situations that they had in the previous week; the other subjects had the order different from the one they had the previous week.

Results

The subjects responded with shorter reaction times under the competitive condition than under the noncompetitive condition. An analysis of variance on these data, equivalent to the one reported for Experiment I, demonstrated that the reaction time was lower under the competitive condition than under the normal condition ($F_{1,18} = 50.4$, $p < .01$).

Figure 2 shows the mean transformed reaction time as a function of trials. Subjects in both groups began with short latency responses and numerous errors, but within 20 trials they reached a latency consistent with few errors. Inspection of the figure suggests that there was an immediate decrease in the reaction time in both groups as soon as the competitive condition was begun.

Using the Kolmogorov-Smirnov statistic as previously described, 5 of the 18 subjects were significantly faster at the .01 level of confidence and 4 of the 18 subjects were significantly faster at the .05 level of confidence in the competitive situation than in the normal situation. No subject was significantly faster (even at the .05 level of confidence) in the normal situation than in the competitive situation.

There was no suggestion that individuals who showed competitive facilitation of the simple reaction time were more likely than others to show competitive facilitation of the complex reaction time. Of the 8 subjects with significant ($p < .01$) facilitation of the simple reaction time, 2 also showed significant facilitation of the complex reaction time; of the 10 subjects who did not show significant facilitation of the simple reaction time, 3 showed significant facilitation of the complex reaction time. The probability of such a result occurring by chance is much greater than .05, and the same lack of relationship was also found using various alternative criteria of competitive facilitation.

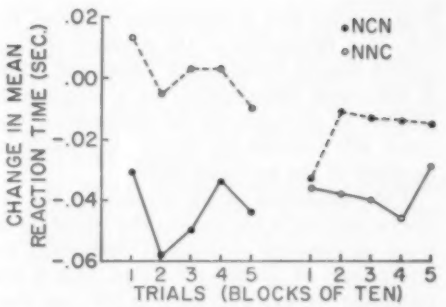

FIG. 2. Change in the mean discrimination reaction time as a function of blocks of 10 trials under the normal and competitive conditions—Experiment II. (Solid lines indicate the competitive condition; broken lines indicate the normal condition.)

Considering a response to the incorrect toggle switch as an error, the number of errors was significantly greater during the competitive condition than during the second normal condition. During the second block of 50 noncompetitive trials, there were 1.8 errors per subject; during the 50 competitive trials, there were 3.0 errors per subject ($p < .01$, sign test). It should be noted, however, that the errors were infrequent under both conditions and that the difference between them was not large. There was no relationship between the amount of increase in the number of errors and the amount of decrease in reaction time during the competitive situation.

Discussion

In a simple reaction time task, the competitive situation resulted in a decrease in response latency; in a discrimination reaction time task, the competitive situation resulted in a decrease in response latency and an increase in errors. These results agree with a long-standing hypothesis

> Rivalry, like social facilitation, increases the quantity, but does not improve the quality, of the output. There is likely in fact, to be a deterioration in quality (Allport, 1924, p. 280).

It is possible that a competitive situation results in a facilitation of the response characteristic that is differentially reinforced. Thus if the speed characteristic of a response is selected as the basis for the competition, the speed will increase and errors may result as an effect of the increased speed.

The explanation of the decrease in the response latency in the competitive situation of these experiments in terms of differential reinforcement of the faster reaction times is not promising. It was undoubtedly the case that the median of the winning reaction times was substantially less than the median of the losing reaction times for each subject. If this differential reinforcement affected the response latency, it would have done so in a gradual manner, and it certainly could not have affected the first trial under the competitive conditions. In the first two experiments, however, there was no evidence for a gradual decrease in reaction time during the competitive trials, and the decrease was as great on the first trial of competition as on any later trial. Of course, any improvement could have been masked by some other factor, such as fatigue, but there was no improvement in reaction time relative to the change found in the normal reaction time condition. Although the differential reinforcement effect might appear with substantially more trials, in the present experiment differential reinforcement cannot account for the competitive facilitation in the reaction time.

To account for the abrupt decrease in reaction time at the introduction of a competitive condition, some assumptions must be made regarding the past experience of the subjects. One possibility is that in the competitive situation subjects' level of motivation was increased and this increase was reflected in their performance. There was reason to believe that the facilitation observed in the previous two experiments was motivational in origin. Subjects appeared to be bored during the noncompetitive trials, but many of them appeared to be more alert during the competitive trials. In the simple reaction time experiment, particularly, many subjects reported that they were bored or sleepy, and they spoke of the monotony of the light and the tone. The comments made at the start of competition to each other and to the experimenter, indicated a heightened level of interest. Of course, further work is essential to measure the motivational changes under a competitive situation, and to determine whether they are necessary for the increase in response speed under the competitive situation.

EXPERIMENT III: MOTIVATIONAL FACTORS IN THE COMPETITIVE FACILITATION OF THE SIMPLE REACTION TIME

Three types of measures of motivational factors involved in competitive facilitation of the reaction time were used in the third experiment:

1. The level of palmar skin conductance was taken during the experiment as a measure of one of the peripheral physiological indices of the level of activation.

2. A self-rating scale regarding the degree of alertness of the subject was used to determine the subjective level of activation.

3. Finally, the typical paradigm of the performance vs. learning experiments was used, in which the level of performance in extinction is related to the level of drive during training (Spence, 1956). In this case, the speed of reaction during the third block of trials was related to the condition, competitive or noncompetitive, during the second block of trials. Many motivational variables are effective when present, but do not have any continuing effect after they are removed. Thus if competition acted like these motivational variables, the condition of the second block of trials would not influence the reaction time during the third block of trials.

The following experiment, then, served to replicate the previous results on competitive facilitation of the reaction time, and to investigate the degree to which motivational variables may be responsible for this effect.

Method

Subjects. Subjects were 56 college students without previous experience with a reaction time apparatus. They were volunteers from small sections of an introductory psychology course and were paid $1.00 for participation in a single 50-minute session. There were 16 pairs of men and 12 pairs of women.

Apparatus. The reaction time apparatus was the same that was used in the first experiment; the

skin conductance apparatus was previously described by Kling, Williams, and Schlosberg (1959), except for the electrodes. In the present experiment silver palmar electrodes were used. They were circular, about 1 inch in diameter, and they were pressed tightly to the subject's palms by means of elastic. The electrodes were lightly covered by a paste composed of equal parts of talc and glycerin, thinned to the consistency of putty by Ringer's solution (Woodworth & Schlosberg, 1954, p. 140).

Procedure. The procedure was the same as that employed in the first experiment, except for the additional instructions regarding the palmar electrodes. The experimenter said

> I will be interested in two things, the speed of your reaction time and the level of your skin conductance. To measure the skin conductance, I will need to put these bracelets on your palms.

The experimenter would then place the electrodes on the subject's palms, and he would postpone answering questions regarding the meaning of skin conductance until the end of the experiment. The measure of the level of skin conductance was taken after the instructions were given for each block of 50 trials; after the two rest periods, and after every fifth trial. At the end of the experimental session, subjects rated themselves on a seven-point scale in answer to the question

> How alert were you (*a*) during the first 50 trials, (*b*) during the second 50 trials, (*c*) during the third 50 trials.

The end-points of 1 and 7 were labeled Weak and Strong, respectively, and the middle point 4 was labeled Moderate.

Design. As in the first experiment, each subject had 150 trials of the simple reaction time, divided into three blocks of 50 trials. A given block would either be under the normal (N) condition or under the competitive (C) condition. Seven pairs of subjects were randomly assigned to each of the following four orders of conditions; NCN, NNC, NCC, and NNN.

Results

Decrease in reaction time during competition. The two groups of the present experiment that were identical to the two groups of Experiment I (NCN and NNC) were analyzed in the same manner as those of Experiment I to determine whether those results would be replicated. No contradictions between the results of the analysis of Experiment I and the equivalent two groups of Experiment III were discovered. Using the Kolmogorov-Smirnov statistic as previously described, 18 of the 28 subjects were significantly faster (at the .01 level of confidence) in the competitive situation than in the normal situation. No subject was significantly faster (even at the .05 level of confidence) in the normal situation than in the competitive situation. The successful replication suggests that the measurement of the conductance level had no substantial influence on the reaction time performance.

The present experimental design permitted the analysis of the effect of the treatment during the second block of 50 trials on the performance during the third block of 50 trials. The mean reaction time of each subject during the first 50 trials was subtracted from his mean reaction time during the third block of 50 trials. An analysis of variance was performed on these data, a 2×2 factorial experiment with 14 observations per cell. The columns represented the treatment during the third block of trials and the rows represented the treatment during the second block of trials. There was a significant column effect ($F_{1,52} = 41.4$, $p < .01$) which is further evidence that the competitive situation resulted in greater decrease of the reaction time than the normal situation. The interaction effect was significant ($F_{1,52} = 9.4$, $p < .01$), indicating that the amount of competitive facilitation was greater after a second block of normal trials than after a second block of competitive trials. Finally, the row effect did not

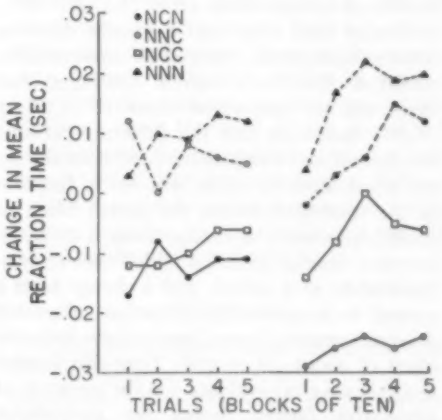

FIG. 3. Change in the mean simple reaction time as a function of blocks of 10 trials under the normal and competitive conditions—Experiment III. (Solid lines indicate the competitive condition; broken lines indicate the normal condition.)

approach statistical significance ($F_{1,52} = 1.5$), indicating that the treatment during the second block of trials did not have an over-all effect on the response latency during the third block of trials.

If subjects learned to respond with short latency responses during the competitive trials, there would be only a gradual increase in the latency when the subjects were returned to the normal situation. Inspection of Figures 1 and 3 suggests that this was the case. The change from the normal situation to the competitive situation produced a large decrease in the mean reaction time, but the change from the competitive situation to the normal situation produced only a small increase in the mean reaction time.

Of course, the analysis of variance gave no support for the involvement of associative factors in the competitive facilitation, but this failure could have been a result of considering the entire third block of trials rather than the initial few trials of the block. Therefore, the analysis of the previous paragraph was repeated, using the mean of the first 10 trials of the third block of trials rather than the mean of the entire block. There was a significant row effect ($F = 19.5$, $p < .01$), indicating that there was a greater over-all reaction time decrease from the normal condition than from the competitive condition; there was a significant column effect ($F = 23.5$, $p < .01$), indicating that there was a greater decrease when subjects were changed to a competitive condition than to a normal condition; but there was no interaction effect ($F > 1.0$, $p > .50$), indicating that the difference between the normal and competitive conditions during the third block of trials was not a function of the treatment during the second block of trials. Apparently a change from a competitive to a normal situation added one- or two-hundredths of a second, and a change from a normal to a competitive situation subtracted an equal amount from a spontaneous recovery effect of about .08 second. Thus this further analysis gave no evidence for the presence of associative factors during the competitive trials (see Table 1).

TABLE 1

CHANGE IN MEAN REACTION TIME FROM LAST TEN TRIALS OF SECOND BLOCK OF TRIALS TO FIRST TEN TRIALS OF THIRD BLOCK OF TRIALS

Second block of trials	Third block of trials	
	Normal	Competitive
Normal	−0.8	−3.4 (−2.3)[a]
Competitive	+0.9 (+0.3)[a]	−0.8

[a] Data from Experiment I.

Increase in conductance during competition. Figure 4 shows the mean level of skin conductance as a function of blocks of five trials for subjects in the four groups. Individual differences in the level of skin conductance were large and reliable. The median correlation between the mean level of conductance during the first block of trials under the normal condition and later blocks of trials under the normal condition was $r = .92$; the median correlation between the mean level of conductance during the first block of trials under the normal condition and later blocks of trials under the competitive condition was $r = .91$. To reduce this source of variance, the mean level of skin conductance of each subject during the first 50 trials was subtracted from his mean level during the third block of 50 trials. An analysis of variance on these data, a 2×2 factorial experiment with 14 observations per cell, demonstrated that the skin conductance was higher in the competitive condition than in the normal condition ($F_{1,52} = 10.5$, $p < .01$). The effect of the

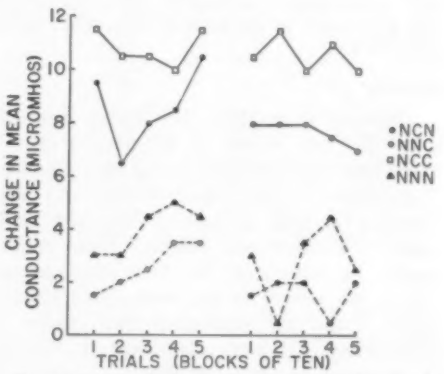

FIG. 4. Change in the mean level of palmar skin conductance as a function of blocks of 10 trials under the normal and competitive conditions—Experiment III. (Solid lines indicate the competitive condition; broken lines indicate the normal condition.)

treatment during the second block of trials on the skin conductance was insignificant, and there was no significant interaction effect.

An inspection of Figures 3 and 4 indicates that the level of skin conductance typically increased when the mean reaction time decreased. No relationship, however, could be discovered between the increase in skin conductance and the decrease in reaction time in the competitive situation. Subjects were divided into those with above median and below median facilitation in each of the measures, "facilitation" defined as the difference between the mean of the first block of trials under the normal condition and the first block of trials under the competitive condition. There were approximately the same number of subjects in each of the four cells, resulting in an insignificant χ^2. The correlation between the amount of increase in skin conductance and the amount of decrease in reaction time was insignificant and in the unexpected direction ($r = -.13$). Although these data were reanalyzed several times using various alternative measures of change of reaction time and of skin conductance, such as the autonomic lability transformation (Lacey, 1956), there was no indication of a relationship between skin conductance and reaction time in this experiment.

Self-rating of "alertness" during competition. The mean rating of alertness was 3.4 during the normal situation and 2.8 during the competitive situation. (The maximum alertness was 1, moderate was 4, and minimum was 7.) Considering the 28 subjects who had both the normal and competitive situations during the second and third blocks of trials, 22 rated themselves more alert during the competitive situation and 4 rated themselves more alert during the normal situation ($p < .01$, sign test). Subjects were divided into those with above median and below median facilitation in alertness, facilitation defined as the difference between the self-rated alertness in the first block of trials under the normal condition and the first block of trials under the competitive condition. As in the case of palmar skin conductance, there was no relationship between the increase in self-rated alertness in the competitive situation and the decrease in reaction time in the competitive situation.

DISCUSSION

As in the previous experiments of this series, Experiment III provided evidence that a competitive situation may result in a decrease in response latency. In addition, it provided evidence that a competitive situation may result in an increase in the level of palmar conductance and in self-rated alertness. Finally, it suggested that the motivational effects of a competitive situation may be independent of its behavioral effect. Although this result was not anticipated, it was not unprecedented, e.g., there are a number of cases in which a physiological measure and a behavioral measure are correlated with a stimulus situation, but are uncorrelated with each other when the level of the stimulus is partialed out. For example, Lansing, Schwartz, and Lindsley (1959) have demonstrated that the alerted reaction time curve and the alpha blocking curve are similar to one another which may suggest that they are influenced by a common process. On the other hand, it was reported the mean reaction time was not a function of alpha conditions under the nonalert condition, and no mention was made of a relation between alpha conditions and mean reaction time of different trials under the alert conditions. Conservatively, then, this study described a relationship between an independent variable, warning signal, and two dependent variables, alpha blocking and reaction time facilitation. Many of the correlations between peripheral indices of level of activation and measures of performance may disappear when the level of the stimulus is partialed out. Such a partial correlation may be established between a peripheral index of activation and speed of performance in a competitive situation, but this study was not successful in obtaining evidence for it.

SUMMARY

Ninety-two subjects were tested on reaction time tasks under normal and competitive conditions. The speed of the simple reaction and of the discriminative reaction increased under the competitive condition. Although the competitive condition resulted in differential reinforcement for fast responses, this differential reinforcement could not account for the increased speed of response since the

increased speed was as apparent on the first trial as on any later trials. Under the competitive condition, the level of palmar skin conductance increased and self-rated alertness increased, but these measures were not related to the decrease in reaction time. Thus the speed of performance and some measures of the level of motivation both increased in a competitive situation, but there was no evidence in these experiments for a causal relationship between them.

REFERENCES

ALLPORT, F. H. *Social psychology*. New York: Houghton Mifflin, 1924.

CHURCH, R. M. Effects of a competitive situation on the speed of response. *J. comp. physiol. Psychol.*, 1961, **54**, 162–166.

DASHIELL, J. An experimental analysis of some group effects. *J. abnorm. soc. Psychol.*, 1930, **25**, 190–199.

FERSTER, C., & SKINNER, B. *Schedules of reinforcement*. New York: Appleton-Century-Crofts, 1957.

KLING, J., WILLIAMS, JOANNA, & SCHLOSBERG, H. Patterns of skin conductance during rotary pursuit. *Percept. mot. Skills*, 1959, **9**, 303–312.

LACEY, J. The evaluation of autonomic responses: Toward a general solution. *Ann. N. Y. Acad. Sci.*, 1956, **67**, 123–164.

LANSING, R., SCHWARTZ, E., & LINDSLEY, D. Reaction time and EEG activation under alerted and nonalerted conditions. *J. exp. Psychol.*, 1959, **58**, 1–7.

SPENCE, K. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.

TRIPLETT, N. The dynamogenic factors in pacemaking and competition. *Amer. J. Psychol.*, 1897, **9**, 507–533.

WOODWORTH, R., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.

(Received May 3, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 41–47

A RELATIONSHIP BETWEEN THREAT IN THE MANIFEST CONTENT OF DREAMS AND ACTIVE–PASSIVE BEHAVIOR IN PSYCHOTICS[1]

JAMES L. FRAMO, JERRY OSTERWEIL,[2] AND IVAN BOSZORMENYI-NAGY

Eastern Pennsylvania Psychiatric Institute

Among the psychoanalytic concepts which have been emphasized as being most central in the explanation of the psychotic process are loss of object relationships (Fairbairn, 1952; Freud, 1958; Guntrip, 1952; Szasz, 1957); tenuous control of impulses, both aggressive and libidinal (Bak, 1954; Hartmann, 1953; Nunberg, 1955, pp. 221–225; Wexler, 1951); defective ego boundaries (Federn, 1952); and overwhelming anxiety (Sullivan, 1947). Notwithstanding the variety of dynamic explanations, in the acute stage of illness before any restitutional structure emerges, the patient experiences amorphous, unbound fears. The concept "threat" as used in this study represents the felt danger of this primitive, unpleasantly affective, preconceptual state. Since this state is so unbearable the psychotic is apt to resort to some unique formulation to account for the whole range of his experiences. The organization and meaning that the psychotic gives to threat, and his behavioral reactions to his attempts to bind and contain threat, constitute the focus of investigation in this research.

Many of the reorganizing efforts of the patient have a transitory, emergency quality while others develop into enduring trends and patterns underlying the protean psychotic manifestations. The study reported here was designed to investigate (*a*) whether two relatively lasting reactive attitudes to threat in psychotics could be delineated operationally, (*b*) whether these inner attitudes could be reliably inferred from specific aspects of overt behavior and the manifest content of the dream, and (*c*) whether a specific lawful relationship between these aspects of behavior and dreams could be established.

[1] A version of this paper was presented at the American Psychological Association Meetings, Cincinnati, Ohio, 1959.

[2] Presently at Student Health Service, University of Pennsylvania, Philadelphia, Pennsylvania.

Reactive attitudes to threat can have many formal properties. The particular formal dimension focused on in this research is the directional or vector quality of threat, whether the self is the target or source of the threat. It is postulated that over a period of time the psychotic tends toward one of two major defensive outcomes in his struggle with threat. In one outcome the patient views threat as originating outside of himself and his outlook toward the world is that he is among enemies; he then acts as if he has to fight for his survival. In the other type of outcome the patient reacts as though the threat were inside himself; he behaves as though there would be dangerous consequences to any action on his part, and consequently shows some variant of inhibited behavior. We have designated the two modal distortions of overt behavior which eventuate from the reactive attitudes of the psychotic, pathologically *active* and *passive*. It is believed that this dimension of psychotic behavior cuts across traditional diagnostic categories and may represent a meaningful classification of psychosis beyond the immediate goals of this study.

It is suggested that the formulation of ascribing threat to others or oneself might be inferred from a variety of psychic productions—delusions, hallucinations, projective test material, and the manifest content of dreams. There has been some controversy over the usefulness of the manifest content of dreams in making inferences about behavior or other aspects of personality. The traditional psychoanalytic viewpoint has stressed the deceptiveness of relying on the manifest content without taking into account the intricate process of dream work via associations leading back to the latent content. One should not, according to this point of view, substitute the associations of the observer for that of the patient (Fleiss, 1953; Freud, 1953;

Kanser, 1954). Another group of workers, among them Alexander and Wilson (1935), Erikson (1954), French (1952), Hall (1953b), and Saul and Sheppard (1956) have demonstrated some regularity in the relationships between manifest content of dreams and personality dynamics. These and most other researchers have worked primarily with the dreams of normals and neurotics, but there has been much less systematic work reported on the dreams of the psychotic (Kant, 1942; Noble, 1951).

In this study it was postulated that the directional aspect of threat in the manifest dreams—whether the self is the target or source of the threat—is a counterpart of the behavioral defensive outcomes. If the patient "fights" the world as though being attacked, it is suggested that he is operating from some threat to himself—a self-directed threat likely to appear in his dreams. If his behavior is seclusive and inhibited, it is suggested that threat toward others would be more prominent in his dreams. It is hypothesized that, in psychosis, there is a complementary relationship between direction of threat in dreams and active-passive behavior. The suggestion for this hypothesis originated from an empirically developed clinical intuition of one of the authors (IB-N) which will be further described in a subsequent paper. Operationally, the hypothesis states: if threat in the manifest content of the psychotic's dreams is directed toward the self, the overt behavior of the patient will be largely characterized by overactivity, and if threat is directed toward others in the dreams, the overt behavior will be largely characterized by pathological passivity.

METHOD

Subjects

Attempts were made to obtain dreams from every available adult patient of the resident population of the Eastern Pennsylvania Psychiatric Institute. Out of a total census of 125 patients, 27 were eliminated for the following reasons: home on leave of absence, neurological involvement, electric shock treatments within the previous month, and subjects with nonpsychotic diagnoses. The latter were eliminated because of the pertinence of the hypothesis for psychotic states. This procedure yielded 98 psychotic subjects who were interviewed for dreams; 92 of these subjects were able to communicate at least one dream.

Although the EPPI is a state research and training psychiatric institution, the distribution of patient characteristics is roughly comparable to that of the "good" wards of the typical state hospital. The emphasis is on relatively short hospitalization, and the patients are largely in an acute phase of disturbance. This group of psychotic subjects had the variety of diagnoses one would expect on an institution-wide basis. The distribution had the following rough breakdown: about half of the subjects were diagnosed as paranoid schizophrenia, about one-fourth as catatonic, hebephrenic, simple, and undifferentiated schizophrenia, and about one-fourth were given a variety of diagnoses such as schizo-affective, manic-depressive, involutional psychosis, and paranoid state. The age range of the subjects was 15–59, with a median age of 32; the median educational level for all subjects was the tenth grade. Fifty subjects were female and 42 were male; 14 of the subjects were Negro. Patients had been hospitalized for periods ranging from several days to 3 years; the median length of hospitalization was 4 months. The sample was relatively heterogeneous with respect to socioeconomic factors. Almost all of the subjects were receiving some form of psychotherapy, individual or group, as well as tranquilizing medication.

Procedure

The purposes of this study dictated that the two sets of data be obtained concurrently and independently so that the dream material could be related to the overt behavior of the patient. A description follows of the procedure utilized for obtaining each set of data, the behavioral judgments and the dreams, and the scoring of the dreams by the Dream Threat Scale.

Active-Passive Behavioral Rankings

Despite the wide variety of symptomatic behaviors which psychotic patients manifest, two common denominators, two modal styles, were postulated which would make themselves felt through their social stimulus value, i.e., the effect they would have on other people. These styles were designated as Active and Passive. Inasmuch as the problems of management of a ward of psychotic patients center on the handling of pathological activity and passivity, it was believed that nurses and aides, who spent most of the time with the patient in the living situation of the ward, would be in the best position to make the judgments. Rather than have these judges make discriminations on an abstract basis the task was constructed in terms that would be meaningful for them. A series of statements of Active and Passive behavior were constructed, largely descriptive phrases of under- and oversocial interaction, although some of the terms were motoric and energy level statements. There was little interest in these or other possible particular descriptive phrases, but rather with the totality of the overall social impact which the patient had on ward personnel. Each cluster of phrases was intended to describe a mode of patient

behavior which would provoke a corresponding remedial or coping attitude from the nursing personnel.

The nursing personnel, who were unaware of the purposes of the research, were seen individually on each of the 10 wards and handed a card with the following phrases on it:

A
Overactive
Bothers other people
Fights or attacks others
Argumentative
Temper tantrums
Moves around too much
Force themselves on other people
Too much energy

P
Too quiet
Others have to approach them
Scared, afraid
Sluggish, slow movements
Overly submissive
Indifferent, apathetic
Withdrawn, isolated
Shy—doesn't mix with others

The following instructions were then read to each *O:*

> This is a list of descriptions about two different or opposite types of patients—what we call the "A" patient and the "P" patient. I am going to read these descriptions out loud and I think you will recognize that the "A" kind of patient makes you control and restrain them, and the "P" kind of patient makes you urge them, help them do things, draw them out.

Each *O* ranked all the subjects on the ward from the most active patient to the most passive. The maximum number of subjects ranked on any ward was 14 and on eight of the wards 10 subjects were ranked.

Rankings were obtained from six nurses and aides on each ward. All six rankings for each subject on each particular ward were than summed, and all the subjects on each ward were then placed in rank order. Subjects above the median of the pooled rank order on each ward were labeled Active, and subjects below the median were labeled Passive. Reliability was calculated for all 10 sets of rankings by means of Kendall's coefficient of concordance. The *W*s ranged between .70 and .81, all significant beyond the .001 level, suggesting that the judges were applying essentially the same standard in ranking the subjects under study, and agreed on the relative rankings.

Inasmuch as each subject was being compared with his peers on his ward in determining his relative ranking along the continuum of activity-passivity, it is possible that a subject labeled Active on one ward would be judged Passive on another. For the majority of subjects the differences between Active and Passive were great enough so that their relative position would probably remain the same no matter what ward they were on. The middle, borderline group would probably have some subjects who would shift across the median. Further, except for the male-female division, incoming patients were randomly assigned to wards, and there was no reason to believe that any one ward contained patients who were behaviorally different from those on other wards.

Dream Collecting

One of the experimenters interviewed the patients and obtained verbatim accounts of all the dreams each subject could remember at a given sitting. Subjects were first asked to tell the dream they had had the previous night. If this was not successful they were asked to report any dream they had had during the previous week, previous month, etc. As a final resort, subjects were asked if they could remember any recurring or outstanding dream from their past life. In addition, subjects were asked to remember any "scary" dreams they may have had. Some of the obtained dreams, then, were of recent origin and others were more distant in time. Old and recent dreams were combined for one of the calculations of this study. The assumption underlying this procedure is that certain conflicts and attitudes persist through life in the unconscious, and since memory is dynamically selective there must be currently pertinent reasons why a *particular* dream from the past happened to be recalled at that particular time.

Although in any dream study one never deals with the direct dream experience but with the reported recollection, with psychotics one could anticipate greater communication difficulties. A few patients confused their dreams with other daytime preoccupations. Some patients were eager to relate their dreams and glad to have the opportunity to talk to someone. Most regarded their dreams as fearful, lifelike realities. The initial response of almost all of the subjects was to disclaim ever having had any dreams; it was almost as if they had been asked if they had ever committed any crimes. A variety of clinical interviewing techniques had to be used to surmount resistances. For example, a variant of the Kinsey interviewing method came to be adopted whereby it was assumed that despite the disclaimer the patient did indeed dream; when the subject would say he never dreams the experimenter would say something like, "Yes, I know but tell me the dream you had last night." Dreams were obtained from all but 6 of the 98 subjects seen. The dreams, which were recorded verbatim in longhand, were then typed up, coded, and submitted to judges for evaluation.

Dream Threat Scale

A special Dream Threat Scale was constructed for evaluating the manifest content of dreams in terms which were pertinent to the study. It is pre-

sented here as it was submitted in the form of instructions to the judges:

You are requested to make several judgments on each of the collection of dreams enclosed with these instructions. Two determinations must be made about each dream: whether or not "threat" appears in the dream, and the direction or target of the threat. Judgments are to be made only on the literal, face value, manifest content of the dream; a sophisticated, clinical attitude about dream work should be inhibited.

I. The first judgment to be made about each dream is whether the dream is a "threat" (T) dream or a "nonthreat" (NT) dream. A "threat" dream, in general, is defined as one in which at *any* point in the dream anyone of the following appears: A dangerous action or image, *or* an affect indicating fear or anxiety, *or* a dangerous situation threatening or potentially deterimental to the survival, health, or way of life of anyone in the dream. A dream should be judged a threat dream under the following conditions:

(a) Where the survival of some person or persons is at stake or where some actual or imminent or completed state of death, injury, or illness occurs to someone in the dream. (Examples: "I dreamt my mother died," or "I saw that my uncle was blind," or "I was being chased by a monster," or " I had to go to the hospital," or "My baby was sick.")

(b) Psychological or "way of life" threats—continued marriage being threatened, financial security, career or job, social ostracism, rejection from or loss of significant relationships, etc. (Examples: "They were brainwashing me," or "I was becoming insane," or "My boss fired me," or "My husband divorced me," or "I told my small son to leave home.")

(c) Fear or anxiety mentioned in dreams with or without appropriate cause. (Examples: "I was frightened"; or "I was walking through the park, the sun was shining, and the children were afraid.")

(d) Natural disasters, fires, explosions, wild animals, or fearful abstractions are considered threatening provided human beings as either the source or the target are being affected in the dream. (Examples: "The city full of people was on fire," or "A tiger was chasing my brother," or "Damnation was on my soul.")

(e) Although there may be no explicitly dangerous content in the dream, threat may be implied by a *potentially* dangerous situation or by some major aspect of one's life situation being potentially altered detrimentally. (Examples: "I was sitting in the car my 4-year-old was driving," or "I was standing on top of a cliff," or "I dreamt my husband was going to Tibet," or "I felt that I may have to look for another job.")

Threat may be determined by the actual or potentially anxiety generating nature of a dream element, relative gravity of several dream elements, or by the dream's emotional context. Resentment, arguing, irritation, annoyance, and other expressions of hostility, as well as words expressing worry, concern, insult, rebuke, frustration, feelings of rejection, and sarcasm *may or may not* be considered threat depending on the significance of the relationships in the dream or the dangerousness of the situation.

II. The second judgment, direction of threat, is to be made only on the "threat" dreams. The whole dream should be summarily evaluated for this purpose.

(a) Threat is directed toward the dreamer *from any source* (Label S for Self).

(b) Threat is directed toward some other person or persons *from any source* (Label O for Other).

(c) The Complex category is used when dreams with more than one dream element have threats going in different directions. If this is judged to be the case it is rated as C for complex with the direction of the most serious or grave threat sublabeled as the dominant one. (For example, Cs would be rated for the following dream: "I was trying to hurt this tiger, jabbing at it with a stick when he turned around and he ran after me and I woke up scared." Example of threat rated as Co: "I felt very sick and weak and I couldn't prevent my baby from dying.")

(d) Threat is involved but direction is unspecified, unable to be determined, or gravity is equally serious to Others and to Self (Label U). (Examples: "There was great destruction all around," or "The building was on fire," or "We were killing each other.")

The reliability of the Dream Threat Scale was determined by the amount of agreement among three trained judges who independently applied these classifications to the total dream sample. The judges were the authors, who had pretested the Dream Threat Scale in a pilot study. Appropriate safeguards were taken to assure that the two sets of variables, dreams and behavior rankings, were evaluated independently. For example, the dreams were judged, without identification, before knowledge of the behavioral rankings was available. Out of a total N of 189 dreams, all three judges agreed on 169 dreams or 89% of the time on whether or not threat was present in the dreams. Insofar as direction of threat is concerned, only the two major categories of Self and Other were dealt with; Self and Complex-Self were combined into Self, and Other and Complex-Other were combined into Other. Using just the threat dreams, the judges agreed on direction of threat 92% of the time.

This Dream Threat Scale, then, was found to be a reliable instrument for making discriminations on such complicated, idiosyncratic data as dreams. It is perhaps partly because of their complexity that few attempts have been made to apply systematic

scoring techniques to the manifest content of dreams (Alexander & Wilson, 1935; Beck & Hurvich, 1959; French, 1952; Hall, 1953b; Harris, 1951; Saul & Sheppard, 1956).

RESULTS

Inasmuch as some subjects contributed more dreams than others (the range was one to four and the modal number was two), it was necessary to establish some criterion which would enable subjects to be independently compared to each other, using one dream for each subject. For the first test of the hypothesis the criterion decided upon was the following: the first dream on which two out of the three judges agreed on direction of threat would characterize that subject's dream behavior. The use of dreams on which all three judges agreed would have reduced the sample too much. Although dreams were obtained from 92 subjects, the total number of patients who gave threat dreams with directions reliably judged was 87.

The hypothesis predicted a corresponding relationship between direction of threat in the manifest content of the dream and Active-Passive behavior, that if threat was Self-directed in the dream the behavior would be characterized by overactivity, and if threat was directed toward Others, by overpassivity.

Table 1 shows that of 52 subjects who gave Self-directed threat dreams, 44 were judged as Active and 8 as Passive; of 35 subjects who gave Other-directed threat dreams, 32 were judged as Passive and 3 as Active. The

TABLE 1

RELATIONSHIP BETWEEN ACTIVE-PASSIVE SUBJECTS AND SELF- OR OTHER-DIRECTION OF THREAT IN DREAMS, USING FIRST DREAM OF EACH SUBJECT ON WHICH A MAJORITY OF JUDGES AGREE

	Self	Other	
Active	44	3	47
Passive	8	32	40
	52	35	87

$\chi^2 = 45.66$, $p < .001$, $df = 1$, Yates' correction applied.

Note.—Although dreams were obtained from 92 subjects, the total number of subjects who gave threat dreams with directions reliably judged was 87. Four Passive patients and one Active patient gave Nonthreat dreams only, which were not used in the calculations.

TABLE 2

RELATIONSHIP BETWEEN ACTIVE-PASSIVE SUBJECTS AND SELF- OR OTHER-DIRECTION OF THREAT IN DREAMS, USING MAJORITY DIRECTION OF THREAT FOR ALL DREAMS OF ALL SUBJECTS

	Self	Other	
Active	37	4	41
Passive	8	22	30
	45	26	71

$\chi^2 = 27.49$, $p < .001$, $df = 1$, Yates' correction applied.

hypothesis was strongly supported by these results.

In this first test all of the data were not used; i.e., just one dream of each patient was used in the calculation and in all but three cases this dream was a recent one. As a further test of the hypothesis it was decided to see whether the predicted relationship would stand up when all of the dreams of each subject were utilized including dreams from the past. Because the different subjects contributed varying numbers of dreams it was necessary to give a unitary characteristic to a given subject's dreams. This was done by establishing a criterion of simple majority —i.e., if a subject gave one Self-directed threat dream and one Other-directed threat dream, that subject was eliminated from the study; the proportion had to be at least two to one in order for a subject to be included. As it turned out, the overwhelming proportion of subjects (71 out of 87) gave dreams with only one type of threat in them. This strong tendency for subjects to dream threat in one direction only, suggests that an enduring feature of the personality may be tapped by this method of analyzing dreams. Apparently even when dreams are remembered from the distant past they tend to be selectively recalled in a way that is lawfully related to the present. The use of this method reduced the N to 71. Of the 16 subjects who gave two threat dreams, one of which was Self-directed and the other, Other-directed, 10 were judged Passive and 6 were judged Active.

The chi square obtained from Table 2 is also highly significant, indicating that when all the data are used and each subject's direction-of-threat-majority is compared to the

behavioral rankings, the predicted relationships are supported.

Discussion

The generalizability of the findings is somewhat limited by the fact that most of the subjects were in an acute phase of psychosis; it is therefore possible that the postulated relationship would not hold up for more chronically hospitalized psychotics. In order for these methods to be applied to normal and neurotic populations, other criteria for Active and Passive behavior would have to be developed. It is questionable, further, that normals and neurotics would report threat dreams as frequently as do psychotics. In our judgment the phenomenon reported in this study applies peculiarly to the psychotic state where emergency measures result in more extreme reactions.

What, then, are the theoretical implications of these results? Although varied explanations could be offered, the following rationale is suggested: When confronted with ambiguous threat or unbound anxiety one characteristically gives it some sort of meaning so that it can be bound and coped with. The psychotic experiences undifferentiated anxiety as a threat to his very survival and thus is more urgently motivated to find some "personal formula" which organizes and tends to reduce the painfulness of the inner experience. We hypothesize that the quasistable resolution which eventuates from the struggle takes the form of an implicit, polarized formulation: that danger or threat comes from either others or oneself. Once the formulation is reached, a corresponding behavioral reaction pattern is likely to develop. In other words, the psychotic acts as if he had the explanation for his difficulties and dreams his formulation of them. While one patient may "reason," "I am cornered and under attack and I must defend myself," another patient may "reason," "There is something about me which threatens others and I must restrain myself." This formulation, which may be conscious or unconscious, is conceived as being a relatively enduring pattern of organization which underlies seemingly incomprehensible behaviors and productions of the psychotic patient. Although it is not identical with a delusion, it may in certain cases parallel or precede verbalized delusions. Because its logic is affect-determined and autistic it is relatively impervious to reality considerations. We have used this inner formula as a hypothetical construct to explain the systematic covariation of certain aspects of dreams and behavior.

Other investigators have also tended to interpret the manifest content of dreams as reflecting conceptual systems which issue from underlying characterological trends of the individual (French, 1952; Hall, 1953a, 1953b). Conventionally, the manifest content of dreams would be expected to go through day-to-day random variation, depending on the events of the previous day, the status of the drives, and effectiveness of the defenses. Such fluctuating variables might be much more difficult to relate systematically to overt behavior. In this study, however, a formal variable of the manifest content (direction of threat) appeared to be stable enough to be related consistently to overt behavior. It remains an essentially empirical question to determine which aspects of the manifest content are stable (characterological) and which are subject to change arising from situational variants and other factors. The delineation of such stable formal variables may lead to an understanding of the inner conceptual formulations that serve the executive function of guiding behavior so as to reduce unbound anxiety, simplify bewildering complexity, and channelize drives.

Summary

This study suggested that the psychotic, in his attempt to deal with unbound anxiety, developed implicitly one of two polarized formulations—that threat is to be expected from others or from oneself. Active or Passive psychotic behavior was interpreted as an attempt by the patient to counteract a corresponding internal threat which would appear in the manifest content of the dream. It was hypothesized that if threat in the subject's dreams was directed toward the Self, the overt behavior would be largely characterized by overactivity, and if threat was directed toward Others, by pathological passivity.

Two sets of data—the dreams and rankings of behavior—were collected concurrently on

92 psychotic subjects. Six nurses and aides from each of 10 wards of a research and training psychiatric institution ranked the patients on the basis of their Active or Passive behavior, primarily on th basis of each subject's interpersonal impact on them; these rankings were found to have high interjudge reliability. A Dream Threat Scale was specially constructed to analyze the nature and direction of threat in the dreams. Three judges used this scale in evaluating the dreams and showed acceptable interjudge reliability.

The hypothesis was strongly supported—viz., that there is a systematic, complementary relationship between the two variables. It was found that when threat was directed toward the Self in the dream, the subject's behavior was significantly more likely to be judged overactive. Other-directed threat in dreams occurred significantly more frequently among overly passive psychotic subjects.

REFERENCES

ALEXANDER, F., & WILSON, G. W. Quantitative dream studies: A methodological attempt at a quantitative evaluation of psychoanalytic material. *Psychoanal. Quart.*, 1935, **4**, 371–407.

BAK, R. C. The schizophrenic defense against aggression. *Int. J. Psycho-Anal.*, 1954, **35**, 129–134.

BECK, A. T., & HURVICH, M. S. Psychological correlates of depression: I. Frequency of "masochistic" dream content in a private practice sample. *Psychosom. Med.*, 1959, **21**, 50–55.

ERIKSON, E. H. The dream specimen of psychoanalysis. *J. Amer. Psychoanal. Ass.*, 1954, **2**, 5–56.

FAIRBAIRN, W. R. D. *An object-relations theory of the personality*. New York: Basic Books, 1952.

FEDERN, P. *Ego psychology and the psychoses*. (Ed. by E. Weiss) New York: Basic Books, 1952.

FLIESS, R. *The revival of interest in the dream*. New York: International Univer. Press, 1953.

FRENCH, T. M. *The integration of behavior*. Chicago: Univer. Chicago Press, 1952.

FREUD, S. *The interpretation of dreams*. (Ed. by J. Strachey) New York: Basic Books, 1953.

FREUD, S. Psychoanalytic notes upon an autobiographical account of a case of paranoia. In A. Strachey & J. Strachey (Eds.), *Collected papers*. Vol. 12. London: Hogarth, 1958. Pp. 75–78.

GUNTRIP, H. A study of Fairbairn's theory of schizoid reactions. *Brit. J. med. Psychol.*, 1952, **25**, 86–104.

HALL, C. S. A cognitive theory of dreams. *J. gen. Psychol.*, 1953, **49**, 273–282. (a)

HALL, C. S. *The meaning of dreams*. New York: Harper, 1953. (b)

HARRIS, I. D. Characterological significance of the typical anxiety dreams. *Psychiatry*, 1951, **14**, 279–294.

HARTMANN, H. Contributions to the metapsychology of schizophrenia. *Psychoanal. Stud. Child*, 1953, **8**, 177–198.

KANSER, M. A field-theory perspective of psychoanalysis. *J. Amer. Psychoanal. Ass.*, 1954, **2**, 526–534.

KANT, O. Dreams of schizophrenic patients. *J. nerv. ment. Dis.*, 1942, **95**, 335–347.

NOBLE, D. A study of dreams in schizophrenia and allied states. *Amer. J. Psychiat.*, 1951, **107**, 612–616.

NUNBERG, H. *Principles of psychoanalysis*. New York: International Univer. Press, 1955.

SAUL, L., & SHEPPARD, EDITH. An attempt to quantify emotional forces using manifest dreams: A preliminary study. *J. Amer. Psychoanal. Ass.*, 1956, **4**, 486–502.

SULLIVAN, H. S. *Conceptions of modern psychiatry*. Washington, D. C.: William Alanson White Psychiatric Foundation, 1947.

SZASZ, T. S. A contribution to the psychology of schizophrenia. *Arch. Neurol. Psychiat.*, 1957, **77**, 420–436.

WEXLER, M. The structural problem in schizophrenia: Therapeutic implications. *Int. J. Psycho-Anal.*, 1951, **32**, 157–166.

(Received May 15, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 48–52

MOTIVATION OF CHRONIC SCHIZOPHRENICS BY INFORMATION ABOUT THEIR ABILITIES IN A GROUP SITUATION

JAMES C. DIGGORY

University of Pennsylvania

AND ARMIN LOEB

Veterans Administration Hospital, Coatesville, Pennsylvania

Students of the complex class of disorders called "schizophrenia" have shown an increasing interest in the ability of patients to engage in goal directed behavior. Investigators of the differences between "schizophrenic" and "normal" thinking (e.g., Goldstein, 1944; Hanfmann & Kasanin, 1942; Rapaport, 1951) deal with one aspect of the process by which goals are set and achievements attempted. Another group (e.g., Cavanaugh, 1958; Olson, 1958; Rodnick & Garmezy, 1957) tries to discover differences in the motivational arousal of schizophrenics and normals, including the possibility that problems which are specifically *social* in content create more difficulty for schizophrenics. This study belongs in the latter class, though it does not explore the effects of systematic variation in the social or nonsocial content of problems set for the subjects.

Our problem is this: if schizophrenic patients are given differential evidence of their abilities to perform a task, will this affect (*a*) their willingness to volunteer for activities involving the same task and different tasks; (*b*) their awareness of the implications of abilities better or worse than theirs, as reflected in their voting for other patients; and (*c*) their levels of aspiration for different tasks?

Demonstrated or presumed ability to achieve one's goals has been suggested as a component or determinant of dominance feeling (Jack, 1934; Maslow, 1937, 1955; Page, 1936; Updegraff & Keister, 1937) and self-esteem (Diggory & Magaziner, 1959; Diggory & Ostroff, 1962; Diggory, Riley, & Blumenfield, 1960; Diggory & Rothman, 1961). Bellak (1958), drawing on recent developments in ego psychology, notably those of Beres (1956), sums up his attitude about the motivation of schizophrenics: they are deficient in "coping behavior."

Diggory and his associates (Diggory & Magaziner, 1959; Diggory & Ostroff, 1962; Diggory et al., 1960; Diggory & Rothman, 1961) advance the proposition that a person's self-evaluation depends on his estimate of his utility as an instrument for achieving his goals. Self-evaluation is supposed to vary directly with the adequacy of one's abilities for pursuing goals, and with the number and importance of the goals he pursues.

Discussions with schizophrenic patients and reading of case reports suggests that they suffer (among other things) from a sense of failure, inadequacy, and lack of hope that they could better their condition by their own efforts even if they had the opportunity. Of course, with hospitalized patients, this attitude may be due, in part, to lack of encouragement and dearth of opportunities in their environment. Accordingly, in this experiment we (*a*) manipulated evidence given the subjects about their abilities, and (*b*) provided them with opportunities to act on the implications of the evidence. Because of our ignorance of the consequences to the patient of a strong increment in his motivation, the experiment was embedded in a play or recreational context, the abilities allegedly tested were of limited utility, and the opportunities for future action were restricted to the hospital environment of the patients.

The experiment was designed to test the hypothesis that, in schizophrenic patients, evidence of the power of their abilities would lead to an increase in their motivation to use these abilities and that evidence that their abilities were of low power would result either in no increment, or in a decrement of their motivation. As evidence of their motivation,

we asked them to volunteer for activities more or less closely related to the abilities which had been tested. The variety of these activities provided an opportunity to explore the generality of any arousal of motivation. In order to discover whether the patients responded appropriately to evidence of different levels of ability, we also asked them to vote for other patients (whose abilities they knew about) for the same tasks as those for which they had offered themselves as volunteers. We compared patients from different wards, varying in their average degree of pathology in order to see whether level of illness affected either volunteering or voting.

METHOD

Subjects and Procedure

Subjects were 72 male chronic schizophrenics, without known organic impairment, from three continuous treatment wards at a Veterans Administration Hospital. All were privileged, but the three wards differ roughly in the degrees of morbidity of their patient populations. Subjects met in groups of three—one member from each ward. Thirty-two three-man groups were run, each in a single session lasting about a half hour. Of these, 24 groups (72 subjects) are included in the analysis. The other 8 groups were eliminated for the following reasons: 1 because one subject claimed he could not read, 1 because the experimenter handed out the wrong tasks, and 6 because the Medium and Low ability subjects' performances were inconsistent. Because they came from different wards, subjects in each group knew each other only casually, if at all.

In each session, subjects were introduced and seated at separate desks. The experimenter explained that he was a psychologist who had been asked by the Recreation Service to invent some new games so the patients would not get tired of playing bingo all the time; and they were now going to play the games developed so far. Rules and scoring procedures were carefully explained and each subject scored his own performances (which was done with consistently high accuracy). The scores were announced aloud and the experimenter recorded them under each subject's name on a blackboard. Each subject completed four tasks.

Each task had a variant for easy, medium, and hard difficulty levels. The answers for each variant were identical but, without the subjects' knowledge, the difficulty of finding the answers varied. For every subject, all four tasks were either easy, medium, or hard. Thus, when scores on all four tasks were reported, one subject was consistently superior to the other two, one was consistently lower, and the third consistently mediocre. The results justify the assumption that subjects interpreted these differences in terms of ability. In half the groups all subjects worked on mental tasks, the remainder did motor tasks. Task type produced no differences on any dependent variables. Only those groups for which the predicted order of performance occurred on *all four* tasks were used.

TABLE 1

EXAMPLES OF MENTAL TASKS

Category	Tasks			Answer
	Easy	Medium	Hard	
Fruit	b_n_na	_n_na	_n_a	banana
Animals	ch_pm_k	c_pm_k	c_pm_	chipmunk
Vegetables	l_tt_ce	l_t_e	_t_e	lettuce
Birds	can_r_	c_n_r_	c_r_	canary

The mental tasks were 3-minute word games, each a list of 20 words of a particular category: fruits, animals, vegetables, and birds. Every subject had the same words, in the same order, but the number of blanks he had to fill in to complete each word differed. There was only one possible correct answer for each word and the score on each game was the total number of words correctly completed. Table 1 gives an example of each of the mental games.

The subjects received two identical copies of each motor task, with instructions to turn to the second as soon as he had finished the first, which he turned face down to prevent copying. The first motor game required drawing a line through the numbers 1 to 50 consecutively, with the numbers, in irregular arrays, being more or less difficult to locate. The score on the first game was the number of consecutive integers correctly connected. The second motor game required the subject to connect successive letters of the alphabet, and was scored by counting the number of consecutive letters connected. In the third game, the subject had to trace a line through a series of gaps without touching either side of the gap. The gaps varied in width, narrowest in the hard task, widest in the easy. Score was the number of gaps successfully traced through. The fourth motor task required mirror tracing between the borders of a 20-sectioned design without touching either side. The space between the borders was widest for the easy condition, narrowest for the hard. Score was the number of sections correctly traced through. The time for each motor game was 2 minutes.

After the four games were completed, while the scores were still on the blackboard, the subject answered 10 questions, each on a 3″ × 5″ card:

1. His guess, anywhere between 0 and 100, of how many different English words he could write in 3 minutes without his actually expecting to write them.

2. His willingness to volunteer to compete with other groups on the same games just played.

3. His vote for one of the three men, including himself, to compete with other groups in the games just played.

4. His willingness to volunteer to compete with other groups on the same kinds of games, but *harder* ones.

5. His vote for one of the three men, including himself, to compete with other groups on the same kinds of games, but *harder* ones.

6. After the experimenter explained how to play one of the motor games to those who just played the mental games, and vice versa, the subject stated his willingness to volunteer to play the new game with other groups.

7. His vote for one of the three men, including himself, to compete with other groups in the new game.

8. After the experimenter said that some men did much worse on the games than anyone present (a statement made largely to alleviate anxiety in those who performed relatively poorly) the subject expressed his willingness to try and teach them to play the games.

9. His enjoyment of the games played.

10. His willingness to meet again with the same group to play some other games.

RESULTS

The question is whether "ability," as indicated by repeated "success" relative to others, differentially affects attitudes toward the self (volunteering) and toward others (voting for them). The principal results are shown in Table 2. Volunteering to compete with others in the same games was high in all ability conditions and did not differ significantly as a function of manipulated ability. But differential ability did affect volunteering for games of the *same kind, but harder*. The so-called "high ability" subjects did this significantly more often than the "low ability" subjects. When subjects volunteered for a different kind of game no differences occurred as a function of ability level. The incidence of volunteering to teach others is also unrelated to ability level, probably because the "others" were said to have done worse than any of the subjects. To summarize the results on volunteering: the majority of subjects (an average of about 59%) regardless of their own supposed ability, volunteered to compete with others in the same games, to compete with others in different games, and to teach less able people to play the games better. The subject's perception of his own ability makes a difference in the likelihood

TABLE 2

FREQUENCIES OF VOLUNTEERING AND RATING BY EXPECTED TASK TYPE AND INDUCED ABILITY LEVEL

Frequency of Volunteering

Induced ability level	Type of task expected							
	Same		Same but harder		Different		Teach others	
	Yes	No	Yes	No	Yes	No	Yes	No
High	16	8	15	9	14	10	12	12
Medium	14	10	11	13	16	8	15	9
Low	13	11	8	16	14	10	14	9
			χ^2 (High vs. Low) $= 5.26$, $df = 1$, $p < .05$					

Frequency of Votes Received

	Type of task expected					
	1 or 2 votes	None	1 or 2 votes	None	1 or 2 votes	None
High	20	4	21	3	15	9
Medium	10	14	12	12	12	12
Low	8	16	6	18	9	15
	$\chi^2 = 12.7$, $df = 2$, $p < .01$		$\chi^2 = 19.2$, $df = 2$, $p < .001$			

that he will volunteer only when (*a*) there is a challenge, i.e., when the expected task is harder than one he has done already and (*b*) one in which his ability has already been tested to some extent. The subjects of high induced ability volunteered for all tasks with about the same frequency, so in the "same-but-harder" tasks it is the *refusal* of the low ability subjects to volunteer which produces the association with ability level.

Voting for another man in the group to compete with others on the *same* game, or in a game of the *same type, but harder*, was directly related to demonstrated "ability." Voting for others to compete in different games varied in the same direction, but not significantly. Subjects did not vote for each other as teachers. So the pattern of voting generally varied directly with perceived ability of the person voted for, regardless of whether the expected task would be more challenging. Since subjects were explicitly permitted to vote for themselves (a kind of volunteering), it is appropriate to mention briefly that incidence of self-voting, though low throughout, paralleled incidence of volunteering, but the trends were not significant.

Mean level of expectation, indicated by the subject's estimate (between 0 and 100) of how many different words he could write in 3 minutes, was significantly affected by the induced ability levels. The means were 52 for high ability, 45 for medium, and 25 for low ($F = 5.81$, $p < .001$).

When asked if they enjoyed the games, the subjects' responses were mostly affirmative: 23/24 for high ability, 22/24 for medium, and 20/24 for low. The proportions of affirmative responses were lower when it came to the question of meeting with the same group to play other games: 12/24 for high ability, 16/24 for medium, and 15/24 for low ability. Neither stated enjoyment nor desire to join the group again were significantly associated with induced ability level.

The choice of subjects from three different wards permits a rough estimate of the effect of degree of morbidity on volunteering, voting, and level of expectation because the wards were ordered in the average "sickness" of their patients. To be sure, this ordering is not perfect, since assignment of a patient to a particular ward was determined not only by how sick he was but also by other factors, such as the location of available beds. Since there is probably considerable overlap in morbidity among the three wards, the data on the effects of morbidity have a lot of error in them, and the effects are probably minimal.

TABLE 3

VOLUNTEERING AND VOTING BY ADJUSTMENT OF PATIENTS IN DIFFERENT WWADS

Patient adjustment	Volunteering for tasks:					
	Same		Same but harder		Different	
	Yes	No	Yes	No	Yes	No
Better	19	5	15	9	17	7
Worse	24	24	19	29	27	21
	$\chi^2 = 4.80$, $df = 1$, $p < .05$		$\chi^2 = 3.43$ (need $\chi^2 = 3.84$ at $p = .05$)			
	Voting for others to work on tasks:					
	Same		Same but harder		Different	
	1 or 2	None	1 or 2	None	1 or 2	None
Better	12	12	14	10	12	12
Worse	26	22	25	23	25	23

There is no evidence from volunteering, voting, or level of expectation, that presumed morbidity level and manipulated ability level interacted. Furthermore, subjects from the allegedly poorer wards gave no consistent evidence of an ordering of adjustment. So Table 3 shows the data on volunteering and voting without regard to ability level and with the two more "poorly adjusted" groups combined for comparison with the "better" group. The notable effect is that subjects from the better ward volunteered to compete in same tasks, and in same-but-harder tasks, in higher proportions than those from the "poorer" wards. But morbidity is not significantly associated with volunteering for different tasks. Nor is voting significantly affected on any issue by the ward of origin of the patient voted for.

The patients' voting behavior indicates that they are aware of the implications of demonstrated "ability" levels *in others*. Also,

their manipulated ability level generalized systematically to their levels of expected performance in naming words. Thus, their pattern of volunteering, where *their own* ability was at stake, cannot be explained by supposing their failure to comprehend the ability scores. Only in the same-type-but-harder task do the low ability subjects volunteer with less frequency than the high or medium ability subjects. For "same" and "different" tasks and for teaching others, volunteers are in the majority, regardless of ability level. The experiment affords no explanation for this outcome. The most reasonable assumption is that the low ability subjects would have welcomed another chance to demonstrate that their ability was higher than it seemed to be. Thus, the subject might suppose that more experience with the *same* task might allow him to improve, while a *different* task might turn out to be one in which he could do well at the outset. Or it may be that they were simply moved to avoid failure only where it appeared to be certain, i.e., the low scoring subjects facing a task of the same type, but harder.

With the exception of the same-type-but-harder task, it appears that when the *self* is involved, these patients were more assertive than could be predicted from the way they applied the same evidence in voting for *others*.

SUMMARY

Chronic schizophrenic subjects varied their level of expected performance in another task directly with their manipulated perceived ability. Also their rate of volunteering was related to perceived ability only when the expected task was harder but one in which their abilities had already been tested.

By contrast, incidence of voting for others was related to their presumed ability regardless of type of expected task.

Better adjusted patients volunteered more often than "sicker" ones for competitions on the same tasks and same-but-harder tasks, but voting had no association with the presumed "sickness" of the person voted for.

REFERENCES

BELLAK, L. *Schizophrenia: A review of the syndrome*. New York: Logos, 1958.

BERES, D. Ego deviation and the concept of schizophrenia. *Psychoanal. Stud. Child*, 1956, **11**, 164–235.

CAVANAUGH, D. K. Improvement in the performance of schizophrenics on concept formation tasks as a function of motivational change. *J. abnorm. soc. Psychol.*, 1958, **57**, 8–12.

DIGGORY, J. C., & MAGAZINER, D. E. Self-evaluation as a function of instrumentally relevant capacities. *Bull. Ass. Int. Psychol. Appl.*, 1959, **8**, 2–19.

DIGGORY, J. C., & OSTROFF, B. Estimated probability of success as a function of variability in performance. *Amer. J. Psychol.*, 1962, **75**, 94–101.

DIGGORY, J. C., RILEY, E. J., & BLUMENFELD, R. Estimated probability of success for a fixed goal. *Amer. J. Psychol.*, 1960, **73**, 41–55.

DIGGORY, J .C., & ROTHMAN, D. Z. Values destroyed by death. *J. abnorm. soc. Psychol.*, 1961, **63**, 205–210.

GOLDSTEIN, K. Methodological approach to the study of schizophrenic thought disorder. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia*. Berkeley, Calif.: Univer. California Press, 1944. Pp. 17–40.

HANFMANN, E., & KASANIN, J. S. Conceptual thinking in schizophrenia. *Nerv. ment. Dis. Mongr.*, 1942, No. 67.

JACK, L. M. An experimental study of ascendant behavior in preschool children. *U. Ia. Stud. Child Welf.*, 1934, **9**, 7–65.

MASLOW, A. H. Cited in G. Murphy, L. B. Murphy, & T. M. Newcomb (Eds.), *Experimental social psychology*. (2nd ed.) New York: Harper, 1937. Pp. 437–440.

MASLOW, A. H. Comments on Professor McClelland's paper. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1955*. Lincoln: Univer. Nebraska Press, 1955. Pp. 65–69.

OLSON, G. W. Failure and the subsequent performance of schizophrenics. *J. abnorm. soc. Psychol.*, 1958, **57**, 310–314.

PAGE, M. L. The modification of ascendant behavior in preschool children. *U. Ia. Stud. Child Welf.*, 1936, **12**, 1–69.

RAPAPORT, D. *Organization and pathology of thought*. New York: Columbia Univer. Press, 1951.

RODNICK, E., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln, Nebr.: Univer. Nebraska Press, 1957.

STOTLAND, E., THORLEY, S., THOMAS, E., COHEN, A. R., & ZANDER, A. The effects of group expectations and self-esteem upon self-evaluation. *J. abnorm. soc. Psychol.*, 1957, **54**, 55–63.

STOTLAND, E., & ZANDER, A. Effects of public and private failure upon self-evaluation. *J. abnorm. soc. Psychol.*, 1958, **56**, 223–229.

UPDEGRAFF, R., & KEISTER, M. E. A study of children's reactions to failure and an experimental attempt to modify them. *U. Ia. Stud. Child Welf.*, 1937, **13**, No. 4.

(Received May 21, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 53–57

CHANGES IN CHILDREN'S VERBAL BEHAVIOR AS A FUNCTION OF SOCIAL APPROVAL AND MANIFEST ANXIETY

VINTON ROWLEY AND E. DUWAYNE KELLER[1]

Department of Psychiatry, State University of Iowa

Studies in the general area of conditioning of verbal behavior have been reviewed by Krasner (1958) and in part are included in the reviews of Adams (1957) and Salzinger (1959). It is apparent that numerous experimental investigations recently have been focused on the modification of verbal behavior as a function of verbal or other rewards of a "social approval" nature. Most of the studies reported achieved positive results using verbal and/or movement (head nod) reinforcers. Over half of the studies reported that none of the subjects evidenced awareness and only 5% of all subjects in the 31 studies combined became aware of what was taking place. It is suggested that verbal behavior can be modified through some rather subtle means which may have important implications for the study of interpersonal relations, particularly in areas where verbal interaction plays a significant role; e.g., psychotherapy, public education, child rearing.

Concerning operant conditioning of verbal behavior, Krasner (1958) states that:

> the specific task given to the *S* is *not* presented as a learning one. The *S* is *not* asked to "learn" a given series of words, nonsense syllables, or paired associates, nor is he even told that "learning" is involved in the studies. Rather the *S* is asked to emit verbal behavior in terms of a given task, and the examiner attempts to reinforce a preselected class of his "verbal behavior" by carefully controlled verbal or nonverbal behavioral cues (p. 149).

The majority of these investigations utilize the paradigm of Skinner; i.e., the dependent variable being the verbal behavior of the subject while the independent variable is the reinforcement which is under the control of the experimenter. The largest group of studies used a variation of the "make-up a sentence" technique first reported by Taffel (1955) in which the subject is shown a series of stimulus cards containing a verb and several (four to six) pronouns. The subject is instructed to make up a sentence starting with any one of the pronouns and using the verb.

With few exceptions (Cox & Cox, 1956; Estes, 1945; Patterson, Helper, & Wilcott, 1960), studies in this area have been carried out with adult subjects and none have reported using the Taffel (1955) technique with children. However, questions may arise concerning the generalization of results obtained with adult subjects to children and the feasibility of using this technique with children. Gewirtz and Baer (1958a, 1958b) have contributed significantly to the operant conditioning technique with young children using social reinforcement. However, it seems worthwhile to study specifically the modification of children's verbal behavior as a function of the social approval of an adult. Further, it may be useful to investigate the effects of anxiety on changes in verbal behavior since it is generally considered that anxiety affects conditioning and many other types of learning. Whether anxiety is impairing or facilitating, however, depends on the complexity of the task (Castaneda, Palermo, & McCandless, 1956; Farber, 1954; Farber & Spence, 1953; Patterson et al., 1960; Taylor, 1956).

The present study proposed to investigate public school children's increased use of a class of verbal behavior as a function of manifest anxiety and two types of social approval (reinforcement) using the Taffel technique.

METHOD

Subjects

Subjects were 90 public school children from an Iowa town of 8,000. They ranged in age from 9 to 12 with a mean of 10.39 years and were in the

[1] The authors express their sincere appreciation to Floyd Davis, Superintendent of the Knoxville, Iowa, Public Schools; James Comer, Elementary School Principal; and to all the teachers involved for their able assistance and cooperation.

fourth, fifth, and sixth grades. Age and grade were proportional throughout with each of the three age-grade levels making up one-third of the group. There were 60 boys and 30 girls—20 boys and 10 girls from each grade. This is comparable to the sex ratio for emotionally disturbed children seen at the State University of Iowa's Child Psychiatry Services and was established because the present investigation has implications concerning clinical activities and future related research with such children.

All subjects had an IQ of 85 or above according to school records, had no apparent physical incapacity or handicap, were not under narcotic medication, and were not misplaced more than one age-grade level in school. No subject had been a resident of an institution (orphanage, delinquent, or psychiatric) and none had a history of arrests or juvenile court involvement.

Materials

The materials consisted of the Children's Manifest Anxiety (*CMA*) scale (Castaneda, McCandless, & Palermo, 1956) and 60 stimulus cards.

The stimulus cards were $3'' \times 5''$, white, unlined index cards. Sixty commonly used, neutrally toned, past tense verbs at a reading level of third grade or below were selected. A different one was typed in capitals in the middle of each card. Six pronouns—I, WE, HE, SHE, THEY, YOU—were typed in capitals below each verb with their respective orders randomized for the series of 60 cards.

Task

The experimental task consisted of making up sentences using the verb on the stimulus card and the subject's choice of one of the six pronouns.

Procedures

The *CMA* scale was group administered in individual classrooms by the classroom teacher. Standard *CMA* scale instructions were given.

The *CMA* scale was scored according to standard procedure and a high-low median split made (Median = 16.79). The children were placed in sampling groups according to sex, grade, and high-low anxiety, and names were drawn randomly from each to compose the treatment groups. There were three treatment groups:

1. Group VA, in which the experimenter responded immediately with verbal approval, using a smile and the word "Good" said in a moderately enthusiastic and sincere manner, following each sentence in which the subject used the personal pronouns I or WE (class of verbal behavior).

2. Group PMA, in which the experimenter responded immediately with physical movement, using a smile and vertical head nod in an approving manner following sentences using I or WE.

3. Group C, a neutral or control group, in which the experimenter made no response following any sentence.

Each treatment group was composed of 15 high anxious (HA) and 15 low anxious (LA), 5 HA and 5 LA from each of the three grades, with the 2:1 sex ratio maintained throughout. Anxiety (A) Scale and Lie (L) Scale scores ranged from 2 to 35 and 0 to 9, respectively, for the experimental sample. One subject had an L score of 9, two had scores of 8, and all others had L scores of 7 or below.

The principal introduced the investigators to the children as Mr. Rowley and Mr. Keller from the University of Iowa. He told them that names had been drawn to determine participants for a language activity with the experimenters. Subjects were taken to the experimental room individually and seated at a table directly across from the experimenter.

The following instructions were given:

> This is a language activity in which you may make up sentences. Here is a stack of cards with words on like this [the experimenter pointed to stack with sample card on top]. There is a word in the center [the experimenter pointed] of each card. For each card, look at the word in the center, then look the bottom words over, and then choose any *one* of the bottom words to *start* a sentence that uses the word in the center. That is, make up a sentence starting with any *one* of the bottom words and then the one in the middle. It doesn't matter if your sentence is long or short—you make up any sentence you want to. Most people find it helpful to just say the first sentence that pops into their mind. This activity will not affect your school marks or standing in any way. I will use a type of shorthand to write down all the sentences you make up. OK? Any questions? You ask them now because after we get going it will be hard for us to stop the activity to talk about them. Remember then, make up a sentence starting with any *one* of the bottom words and then the word in the middle. Here we go with the first card.

The first card was turned over followed by each in the stack. Subjects did not report, nor did their questions indicate, difficulty in understanding the procedure and none perseverated on any one pronoun.

All subjects were treated alike during the presentation of the first 15 cards to establish a base level. The experimenter exposed the stimulus card, recorded the response, and exposed the next card. This sequence constituted Block 1 and the next 45 cards were divided into Blocks 2, 3, and 4 consisting of 15 cards each. The experimenter responded differentially, according to the treatment groups above, on all cards following Block 1. The cards were thoroughly shuffled between subjects in an attempt to control order effects. The experimenter alternated treatments from subject to subject in verbal approval, physical movement approval, and control sequences to randomize experimenter day-by-day differences throughout treatment groups. An equal number of subjects from each treatment group were run each day. It required 3 consecutive days to run the

90 subjects, 30 each day. About 12–15 minutes were needed for each subject to complete the task.

Following completion of the task, subjects were asked the following questions to give some indication of the awareness of the relation between their responses and the experimenter's responses (Taffel, 1955):

1. Did you usually give the first sentence that entered your mind?
2. How did you go about deciding which of the words on the bottom to use?
3. Which do you think you used the most times?
4. Why?

Two experimenters, the investigators, were employed. They each ran half the boys and girls in each grade-treatment anxiety group. Such a procedure provided some control over possible experimenter differences by exposing the groups equally to the subjects.

A difficulty was encountered during the preliminary testing of the methods which should be noted. To serve as an effective reinforcer, the inaudible head nod must be seen by the subject. However, some preliminary subjects continued to look at the card or elsewhere following their response. A procedure was adopted in which the experimenter did not record the subject's response until he looked up. Within a few trials, all subjects looked at the experimenter following each sentence. The experimenter then recorded the sentence and turned the stimulus card over.

RESULTS

The criterion measure for all subjects was the frequency of sentences beginning with WE or I within each block of 15 trials. The effect of social reinforcement upon the WE and I class of pronouns was to increase the response frequency of this class. The usage of WE or I in the first 15 sentences was selected as the basal level of frequency. An analysis of variance was performed on the Block 1 data to determine if any significant differences could be demonstrated at the outset of training. This analysis yielded nonsignificant findings ($F = .932$).

In order to assess the effects of anxiety level for each treatment condition, a Type III analysis of variance (Lindquist, 1953) was conducted, in which anxiety level and treatments constituted "between-subjects" factors, while trials and those interactive effects which included trials were "within-subjects" factors. The summary table of the results of this analysis of variance is presented in Table 1.

The analysis indicated that neither anxiety nor its interactive effects were significant. The treatment effect was significant ($p <$ .001). It will be noted that the Trials × Treatment interaction is significant ($p < .001$), in which the difference among the treatments vary over trials.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF THE USE OF REINFORCED PRONOUNS AS A FUNCTION OF TREATMENT CONDITIONS AND ANXIETY

Source	*df*	*MS*	*F*
Between subjects	89	30.94	
Treatment Conditions (C)	2	312.10	12.43*
Anxiety (A)	1	10.34	.41
C × A	2	5.27	.21
Error (b)	84	25.10	
Within subjects	270	4.87	
Trials (T)	3	48.52	12.13*
T × C	6	20.52	5.13*
T × A	3	1.76	.44
T × C × A	6	5.39	1.35
Error (w)	252	4.00	
Total	359		

* Significant at .001 level.

Separate Type III analyses were done to assess the effects of sex and examiner upon the rate of acquisition of the reinforced class of pronouns. Sex or examiner influences and their respective interaction effects were not significant by these analyses.

In view of the nonsignificant results with respect to anxiety level, sex of subject and examiner, these groups were collapsed over treatments and analyzed by a Type I analysis of variance (Lindquist, 1953). Graphic representation of these data is presented in Figure 1. Table 2 contains a summary table of the results of this analysis to determine whether the frequency of the criterion measures differed among VA, PMA, and C groups over trials. VA and PMA groups demonstrated significant increases while the C group did not manifest an increased usage of the reinforced class of pronouns.

A Treatments × Trials interaction was significant, indicating that differences among these groups varied over trials. An analysis of variance of Block 4 indicated that these groups differed significantly ($F = 15.43$, $df = 2/87$, $p < .001$). The differences between the three means revealed that both the VA and PMA groups learned significantly more than the controls ($t = 5.47$, $p < .001$ and $t = 2.59$, $p < .02$, respectively). The difference between VA and PMA indicated that

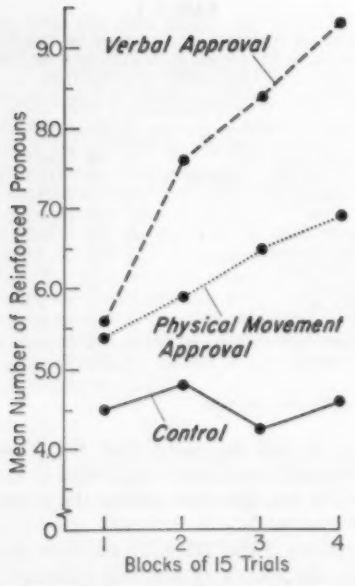

FIG. 1. Mean number of reinforced pronouns used per block of 15 trials.

the verbal reinforcement was more effective than the head nod ($t = 2.33$, $p < .05$). To determine if these three groups used the first person pronouns more frequently on the fourth block than on the first block, a t test for related samples was carried out over these two trial blocks for each treatment group. The results indicated that significant increases in preference for WE and I pronouns occur in the VA group ($t = 6.22$, $p < .001$) and PMA group ($t = 2.06$, $p < .05$), but not in the C group ($t = .26$).

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF THE USE OF REINFORCED PRONOUNS AS A FUNCTION OF TREATMENT CONDITIONS

Source	*df*	*MS*	*F*
Between subjects	89	30.94	
Treatment Conditions (C)	2	312.10	12.75*
Error (b)	87	24.48	
Within subjects	270	4.87	
Trials (T)	3	48.52	12.11*
T × C	6	20.52	5.12
Error (w)	261	4.01	
Total	359		

* Significant at .001 level.

Since the method for assessing awareness in our subjects did not reveal any "aware" subjects, the influence of this factor could not be evaluated in this study.

DISCUSSION

The previous findings of learning without awareness are generally confirmed by the data of the present study. In both treatment groups, significant increased preference of the reinforced pronouns over trials occur while the control subjects failed to show this result. However, neither the influence of different experimenters or sex of subject on the rate of acquisition of these responses could be demonstrated. The method of questioning utilized in the present investigation failed to reveal any subject who was able to state the contingency that the increased usage of the criterion response had occurred in association with the verbalization of approval or physical expression of approval by the experimenter. In view of the lack of any direct evidence of this contingency, it would seem justifiable to assume that learning in the predicted direction occurred without conscious awareness by subjects of the intentions of the experimenter.

Significant learning took place in both treatment groups with the verbal reinforcement being significantly more efficient than the nonverbal reinforcement. This finding substantiates previous results obtained with adult subjects.

The influence of anxiety or motivational level as implied by Castaneda, Palermo, and McCandless (1956) was not demonstrated. The findings of Patterson et al. (1960), using "disturbed" children and a more complex task, support this hypothesis. It should be noted, however, that in the present investigation the task was not presented as a "learning" task per se and the failure of the *CMA* scale to reflect either facilitation or interference with the rate of acquisition may be attributable to this. Nevertheless, in view of the relative simplicity of this task, one might well have predicted a positive relationship between anxiety and rate of acquisition.

The general research implications proposed by Krasner (1958) seemingly can be extrapolated to children. More specifically, it would seem that more experimental investiga-

tion is needed to assess the factor of "awareness," personality attributes, chronological age, varying amounts and types of reinforcement, and differences between emotionally disturbed and normal children. The degree of involvement of a child with significant prestigeful persons, such as parents, teachers, or psychotherapists and the subsequent influence of these people upon the conditioning of verbal behavior remains an open, but interesting research area.

SUMMARY

This study investigated the operant conditioning of a class of verbal behavior as a function of manifest anxiety and of two types of social approval in grade school children. Anxiety levels were inferred from the Children's Manifest Anxiety scale. Social reinforcers consisted of a "head nod" and the verbalization "good." Acquisition of a verbal conditioned response was obtained in both reinforcement groups ($p < .001$) with the verbal reinforcer being significantly better than physical movement. Anxiety effects were not demonstrable. None of the subjects stated the contingency between the reinforcing stimulus and response class subsequent to the acquisition trials.

REFERENCES

ADAMS, J. K. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, **54**, 338–406.

CASTANEDA, A., MCCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, **27**, 317–326.

CASTANEDA, A., PALERMO, D. S., & MCCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Develpm.*, 1956, **27**, 328–332.

COX, F. N., & COX, MARJORIE J. Unconscious reinforcement of a verbal response in children. *Aust. J. Psychol.*, 1956, **8**, 140–151.

ESTES, KATHERINE W. Some aspects of reinforcement upon verbal behavior of children. Unpublished doctoral dissertation, University of Minnesota, 1945.

FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1954*. Lincoln: Univer. Nebraska Press, 1954. Pp. 1-46.

FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.

GEWIRTZ, J., & BAER, D. Deprivation and satiation of social reinforcers as drive conditions. *J. abnorm. soc. Psychol.*, 1958, **57**, 165–172. (a)

GEWIRTZ, J., & BAER, D. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, **56**, 49–56. (b)

KRASNER, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148–170.

LINDQUIST, E. F., *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

PATTERSON, G. R., HELPER, M. E., & WILCOTT, R. C. Anxiety and verbal conditioning in children. *Child Develpm.*, 1960, **31**, 101–108.

SALZINGER, K. Experimental manipulation of verbal behavior: A review. *J. gen. Psychol.*, 1959, **61**, 65–94.

TAFFEL, C. Anxiety and the conditioning of verbal behavior. *J. abnorm. soc. Psychol.*, 1955, **51**, 496–501.

TAYLOR, JANET A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–328.

(Received May 26, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 58–61

CRITIQUE AND NOTES

TEACHING COMMON ASSOCIATIONS TO SCHIZOPHRENICS [1]

ROBERT SOMMER, GWYNNETH WITNEY
University of Alberta
AND HUMPHRY OSMOND
Saskatchewan Hospital

Bleuler (1955), Arieti (1955), and others have stated that a disturbance in association is the primary symptom of schizophrenic thought disorder and "from it can be derived the majority of secondary symptoms" (Bleuler, 1955, p. 355). If methods for improving schizophrenic association can be developed, this should have practical significance for the rehabilitation of schizophrenic patients. Fortunately, there exists a comparatively simple and objective instrument for measuring goodness of association. This is the Word Association Method which was originally used by Francis Galton in 1879 and later modified by other researchers. Many investigators have shown that schizophrenics have less common associations than normals. Despite this well-established finding there has been no systematic investigation of ways to teach schizophrenics to give more common associations.

The present study is one way of doing this, and investigates the utility of techniques developed by Lindsley (1956) and his associates for motivating regressed patients. In the first part of this study we tried to find out whether rewards of cigarettes or praise could induce schizophrenic patients to give more common associations. In the second part an attempt was made to condition alcoholic patients to give *un*common associations.

METHOD

The subjects in the experiment were nine schizophrenic patients and five alcoholic patients in a large mental hospital. The average age of the schizophrenics was 39 years and their last continuous hospitalization averaged 7 years. The average age of the alcoholics was 47 years and they averaged 2 weeks in hospital for the current admission. No patient was known to be mentally deficient or suffering from organic brain pathology. All of the schizophrenic patients had been tested in previous research, although not with the Word Association Method, and had been reasonably cooperative. All patients were identified as smokers by ward nurses.

The Kent-Rosanoff (1910) list of 100 stimulus words was divided into three sections of 33 words each (the hundredth stimulus word was omitted). A counterbalanced experimental design was used whereby some subjects received Words 67–99 on odd-numbered sessions and Words 1–33 in even-numbered sessions, while other patients received the words in a different order. Each subject was first tested with one list using no reinforcement at all. This was intended to provide a baseline of the individual's use of common associations. During subsequent administrations of the test, each patient was rewarded every time he gave a common response. A common response was defined as any one of the three most common responses to each stimulus word in the Kent-Rosanoff Frequency Tables. These tables contain *all* the responses given by 1,000 normal adults to the 100 stimulus words. For example, for the stimulus word TABLE, the three most common responses were: chair, wood, and eat; for DARK, the three most common responses were: light, night, and black. Following the first nonreinforced trial, the next eight trials, which took approximately 2 weeks, used cigarettes as rewards. Every time a patient gave a common association, he received a cigarette. The same procedure was used with both schizophrenic and alcoholic patients. It was not possible to continue the series for the alcoholic patients further because of their brief stay in hospital. However, we felt it would be useful to continue the trials with the schizophrenic patients using a different reinforcement. During the next week when only the schizophrenic patients were tested, whenever one of them gave a common association, the experimenter said in an enthusiastic and encouraging voice "Good," "Fine," or "You're doing very well." During the fourth week of the study, cigarette rewards were used again but this time the patient was explicitly told

> Each time the word you say is one of the words that most people would say, I will give you a cigarette.[2]

[1] This research was supported in part by a grant from the Department of Health and Welfare, Ottawa.

[2] In a previous study (Sommer, Dewar, & Osmond, 1960) it was found that schizophrenic patients were just as aware of other people of what constitutes a

On the final trial, which for the schizophrenics occurred after 4 weeks but for the alcoholics occurred after 2 weeks, each patient was given a completely new list of 33 stimulus words without any reinforcement. The goal of this was to learn whether there would be transfer of training from the reinforced series to a completely new list of words. In summary, the sessions followed this sequence:

For schizophrenic patients:

Trial 1—No reinforcements
Trials 2–9—Cigarettes as rewards
Trials 10–13—Praise as reward
Trials 14–18—Explicit instructions regarding the reinforcement procedure, plus cigarettes
Trial 19—New list of stimulus words, no reinforcements

For alcoholic patients:

Trial 1—No reinforcements
Trials 2–9—Cigarettes as rewards
Trial 10—New list of stimulus words, no reinforcements

The method used in the second part of the study, where alcoholics were conditioned to give *uncommon* associations, will be described later in the paper.

RESULTS

On the first trial when no reinforcements were used, the alcoholics gave significantly more common associations than the schizophrenic patients. The mean number of common associations given by the alcoholics was 17.8 compared with 11.3 given by the schizophrenics. This was also our finding in a previous study (Sommer, Dewar, & Osmond, 1960) and it is encouraging to see it confirmed with new samples. Figure 1 shows that the use of cigarettes as reinforcements produced an immediate and regular rise in the number of common associations for the alcoholics. There was a smaller increase in the number of common associations for the schizophrenic patients, and this was not maintained throughout the series. The scores of each individual alcoholic usually showed a consistent rise over time; e.g., on the first series of words one alcoholic gave 20, 23, 23, 27, and 31 common associations. The schizophrenics, on the other hand, fluctuated markedly from trial to trial and there was no guarantee that an improvement manifest on Trial 5 would be maintained on the next trial. The individual learning curves show a great deal of fluctuation from day to day that cannot be seen in the group learning curves in Figure 1. The difference between the average score on the first (nonreinforced) trial and the average score on the last

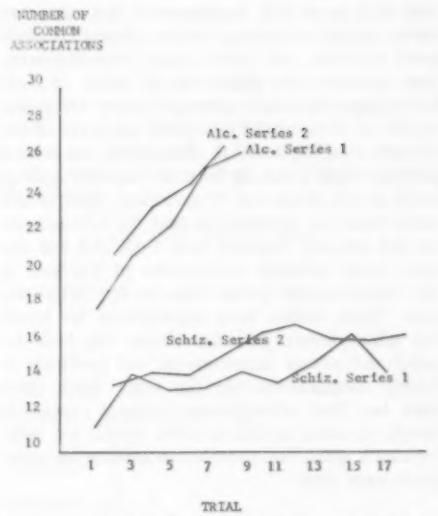

FIG. 1. Rewarding schizophrenics and alcoholics for common associations.

trial of the reinforced series was significant at the .01 level ($t = 5.16$, $df = 4$) in the case of the alcoholics and at the .05 level ($t = 2.99$, $df = 8$) in the case of the schizophrenics.

When the new list of stimulus words was used (Trial 9 for the alcoholics and Trial 19 for the schizophrenics), the alcoholics showed more transfer effect than schizophrenics. From their baseline average of 17.8 common associations, the alcoholics rose to an average of 23.3 common associations on the new list of words ($t = 4.17$, $p < .05$). The schizophrenics, following 4 weeks of reinforcement trials, were only able to raise themselves from a baseline score of 11.3 common associations to an average of 14.5 common associations on the new list of stimulus words ($t = 1.54$, $p = ns$).

DISCUSSION

Since 33 stimulus words were used in each session, this was the absolute ceiling to the number of common responses that a person might give. Because the schizophrenics started out much lower than the alcoholics, they had more room for improvement. Our previous study (Sommer et al., 1960), and also one by Pfenninger cited by Bleuler (1955) showed that schizophrenics vary more than normals in their responses to a second administration of the test. This makes it unlikely that the poor learning of the schizophrenics was due to a greater perseveration of their original responses. The schizophrenics not

common association. In fact the schizophrenics did slightly better than the alcoholics in gauging whether their responses were common or uncommon.

only had room for improvement but also less stable initial responses, which should provide good material for conditioning. Unfortunately, they seemed to be unable to use either of these advantages for their average score after one month of training did not reach the level of the alcoholics *before* training. Essentially the schizophrenics took twice as long to improve half as much as the alcoholics. Nonetheless, there is still some room for optimism in that the schizophrenics *did* actually improve over time. All but one gave more common associations at the end of the reinforcement period than on the initial test trial. These results have implications for teaching schizophrenic patients. Those who look for rapid and steady improvement will probably be keenly disappointed. On the other hand, those who feel that schizophrenic patients cannot be taught to think or talk as other people are overlooking a small but nonetheless genuine improvement over time.

Conditioning of Uncommon Associations

Since the preceding study had shown that some patients can be taught to give common associations to a word association test, we felt it would be useful to see if they could be taught the reverse, i.e., to give *uncommon* responses to stimulus words. This problem has some relevance to creative thinking, since one aspect of creative thought is the use of unusual or uncommon associations. For our purposes we defined an uncommon association as any response to a stimulus word *other than* the three most popular responses on the Kent-Rosanoff Frequency Tables. For example, to the stimulus word TABLE, any response *other than* chair, wood, or eat was considered uncommon.

Five alcoholic patients were used in this experiment, none of whom had been used in the previous study.[3] The procedure was similar to that of the preceding study except that the patients were rewarded for uncommon rather than common responses. Cigarettes were again used as rewards and each subject was tested daily over a 1-week period.

RESULTS

During the one week of reinforcement trials the patients rose from an average of 12.6 uncommon associations to an average of 29.6 uncommon associations. Since the maximum score possible was 33, there seemed little point in continuing the trials beyond this 1-week period. The slope of the curve in Figure 2 leaves no doubt that alcoholic patients can be taught to give uncommon associations.

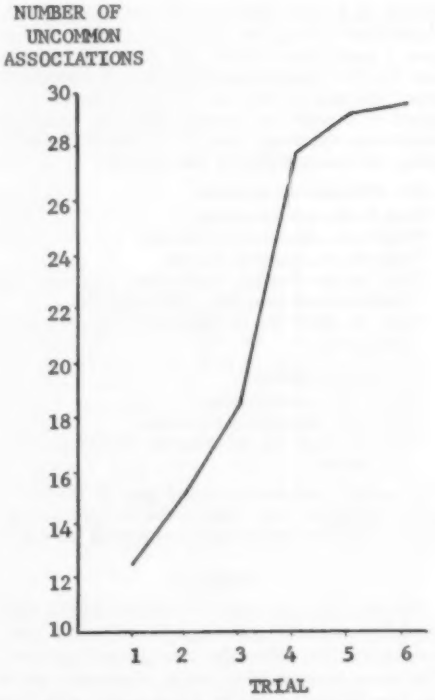

FIG. 2. Rewarding alcoholics for uncommon associations.

SUMMARY

The goal of this study was to learn whether schizophrenic patients could be taught to give common associations on a word association test. Nine schizophrenic patients and five alcoholic patients in a 1,500-bed mental hospital were tested for several weeks using the Kent-Rosanoff Word Association Test. Every time the patient gave a common response he was rewarded with a cigarette or praise. The results showed that the schizophrenic patients improved their performance over time, but not as quickly or markedly as the alcoholic patients. The improvement of the alcoholics was rapid and consistent, while the performance of the schizophrenics was sporadic and inconsistent. In a second experiment five alcoholic patients were rewarded with cigarettes every time a patient gave an *uncommon* association to the Kent-Rosanoff Word Association Test.

[3] There seemed little reason to reward the uncommon associations of schizophrenic patients who already have an excess of them.

In one week's time the patients improved dramatically, almost reaching the maximum possible score on the test.

REFERENCES

ARIETI, S. *Interpretation of schizophrenia.* New York: Robert Brunner, 1955.

BLEULER, E. *Dementia praecox.* New York: International Univer. Press, 1955.

GALTON, F. *Inquiries into human faculty and its development.* London: Dent, 1943.

KENT, GRACE H., & ROSANOFF, A. J. A study of associations in insanity. *Amer. J. Insan.*, 1910, 67 (July, October).

LINDSLEY, O. R. Operant conditioning methods applied to research in chronic schizophrenia. *Psychiat. res. Rep.*, 1956, 5, 118–139.

SOMMER, R., DEWAR, R., & OSMOND, H. Is there a schizophrenic language? *Arch. gen. Psychiat.*, 1960, 3, 665–673.

(Received May 5, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 61–64

OPEN- AND CLOSED-MINDEDNESS AND THE ABILITY TO DIFFERENTIATE SOURCE AND MESSAGE [1]

FREDRIC A. POWELL

Michigan State University

Describing the nature of *dogmatism* in terms of belief-disbelief systems, Rokeach (1960, p. 396) suggested that "open" and "closed" systems basically differ in (*a*) the capacity to distinguish information from source of information, and (*b*) to evaluate information and source of information on their intrinsic merits. In Rokeach's conceptualization, the more closed an individual's belief system, the more difficult it should be for him to distinguish between information received about the world and information received about the source. On the other hand, the more open his belief system, the better should be his ability to evaluate and act upon information about the world, independently and on its own merits.

The purpose of this study was to test, in a more or less realistic communication situation, the validity of this distinction. The study was undertaken in an attempt to determine whether or not open and closed individuals do, in fact, differ in their ability to differentiate between and evaluate independently messages and their sources in a communication context.

Studies of source credibility or prestige (Hovland & Mandell, 1952; Hovland & Weiss, 1951; Kelman & Hovland, 1953) have been concerned with the effect of a source upon the "communicatee's" judgment and acceptance of the ideas and arguments presented in the source's message. In this study, an effort was made to achieve greater generalizability of the findings by treating both the source and the message as "reference points" against which the subjects were to judge new stimuli—sources and statements. It was expected that the influence of open- or closed-mindedness should operate to approximately the same degree and in approximately the same manner regardless of the reference point used by the subjects in judging sources and their messages.

METHOD

Subjects

Seventy-six subjects were drawn from the population of Lansing, Michigan, and the surrounding township. They were selected by means of a modified quota sampling procedure in such a way that the sample was divided equally, on the basis of their subsequently obtained dogmatism scores, between open and closed individuals. The sample was also evenly distributed, as nearly as possible, in terms of age, sex, political party preference, education, and socioeconomic status. Of the open subjects, 16% had 9 years or less of formal schooling, 37% some high school, and 47% had some schooling past high school. Of the closed subjects, 29% had 9 years or less of schooling, 45% had some high school, and 26% had some college. Educational status was not, however, significantly related to dogmatism. The subjects were assigned to one of four groups on

[1] This article is taken from a thesis submitted to the Department of General Communication Arts, Michigan State University, in partial fulfillment of the requirements for the degree of Master of Arts. The author is indebted to Hideya Kumata, P. J. Deutschmann, and L. C. Barrow, Jr., for their guidance and assistance in planning and conducting the study.

TABLE 1

MEAN D^2 SCORES OF THE FOUR GROUPS ON THE SIX SOURCE-MESSAGE PAIRS

Group	Dogmatism/ Reference	Source-Message pairs[a]						Group means
		I	II	III	IV	V	VI	
I	Open/ Source	43.0000	34.1579	34.8421	40.4737	30.9474	34.7895	36.3684
II	Open/ Message	43.5789	39.0000	46.8947	34.5789	36.2632	41.3158	40.2719
III	Closed/ Source	14.2632	13.6842	19.5263	10.2632	10.4737	17.1053	14.2193
IV	Closed/ Message	15.7895	18.7368	24.2105	14.4737	16.9474	20.8421	18.5000
Source-Message pair means		29.1579	26.3947	31.3684	24.9474	23.6578	28.5132	27.3399

[a] The source-message pairs were identified as follows: I: Nixon domestic affairs statement, II: Nixon foreign affairs statement, III: Nixon integration statement, IV: Kennedy domestic affairs statement, V: Kennedy foreign affairs statement, VI: Kennedy integration statement.

the basis of the two independent variables: open- or closed-mindedness, and the reference point used in judging the sources and statements. Each group contained 19 subjects.

Procedure

The subjects indicated on 14 Semantic Differential scales their judgments of the two major Presidential candidates in the 1960 election. Employing the same 14 scales, they also judged three statements made by each of the two candidates in the course of the campaign. The statements of the two sources were matched, as nearly as possible, in their purpose, length, and persuasive appeal. The statements dealt with domestic policy, foreign policy, and racial integration issues.

Group I (open-minded) and Group III (closed-minded) first judged the sources and then judged the statements, attributed to their sources. Group II (open-minded) and Group IV (closed-minded) judged the unattributed statements and subsequently their sources, knowing the statements of the respective sources.

The average source-statement difference or distance (D) in judgment was obtained for each subject. The average distance between source judgments and judgments of the sources' statements was squared, and this score (D^2) taken as the measure of the ability of each subject to differentiate the sources from the content of their statements and to judge each independently (Osgood, Suci, & Tannenbaum, 1957).

In line with the hypothesis, it was expected that the D^2 scores would be significantly greater for the open-minded subjects than for the closed-minded subjects.

The instrument, which was administered in the subject's home, included the sources and statements to be judged, the Semantic Differential scales to be used in judging them, the 40-item (Form E) Rokeach Dogmatism scale, and a series of demographic questions for classificatory purposes.

RESULTS

The mean D^2 scores of each of the four experimental groups (open-minded, source reference; open-minded, message reference; closed-minded, source reference; and closed-minded, message reference) for each of the six source-statement pairs are reported in Table 1.

The individual D^2 scores were summed for each of the four groups for the six source-statement pairs. The data were tested for differences between the summed D^2 scores of the four groups. The procedure described by Collier (1958) for the analysis of variance of correlated measures was employed. The results are reported in Table 2.

The assumption of homogeneity of variance in the data was found to be untenable. The two most extreme group variances (Group I and Group III) were compared with a resultant F value of 1.90 (not significant at the .05 level).[2] For this reason, a correction factor suggested by Greenhouse and Geisser (1959) was introduced. The correction involves reducing the degrees of freedom by a factor $\left(\frac{1}{p-1}\right)$ (where p, in this instance, equals the number of source-statement pairs) for the F tests of the source-statement pairs (F_1) and of the interaction be-

[2] Bartlett's test of homogeneity, although not appropriate to data of this sort, was applied as a means of obtaining another crude measure of the data's heterogeneity. The resultant B value of 172.683 ($p < .001$) indicated the data's extreme heterogeneity.

TABLE 2

ANALYSIS OF VARIANCE OF TOTAL D^2 SCORES OF THE FOUR GROUPS

Source	*df*	*MS*		F[a]
Within individuals				
Pairs	5	617.896	F_1	4.987*
Dogmatism × Pairs	5	280.502	F_{11}	2.268
Reference × Pairs	5	216.344	F_{12}	1.750
Dogmatism × Reference × Pairs	5	157.343	F_{13}	1.272
Individuals × Pairs	360	123.625		
Between individuals				
Open/Closed	1	54978.178	F_{21}	42.513**
Source/Message	1	1908.967	F_{22}	1.476
Interaction	1	4.049	F_{23}	.003
Individuals	72	1293.200	F	

[a] The degrees of freedom for F_1, F_{11}, F_{12}, F_{13} were reduced from 5 and 360 by a Factor 1/5 as suggested by Greenhouse and Geisser (1959) in order to correct for heterogeneity of variance in the data.
* Significant at the .05 level.
** Significant at the .001 level.

tween the pairs and the two independent variables (F_{11}, F_{12}, and F_{13}).

The D^2 scores were found to be significantly greater for the open-minded groups than for the closed-minded groups ($p < .001$).

In addition, it was found that the D^2 scores did not differ significantly between the source reference and message reference conditions. Although the scores in the message reference condition were larger for both the open and closed subjects, this difference, when submitted to Tukey's procedure (Ryan, 1960) for comparing means, was found to be nonsignificant at the .05 level.

No interaction was observed (*a*) between open- and closed-mindedness and the reference point variable, or (*b*) between open- and closed-mindedness and the source-statement pairs.

DISCUSSION

The D^2 scores of the open-minded subjects were found to be significantly greater than those of the closed subjects, providing support for the hypothesis. It is concluded, on the basis of this evidence, that open-minded individuals are better able to distinguish between and evaluate independently the content of a message and the source of the message than are closed individuals. Empirical support was in this way obtained for the validity of the distinction drawn by Rokeach (1960) between open and closed individuals.

It was further observed that the D^2 scores of the source reference and message reference groups were not significantly different, although the scores of the latter groups were larger numerically. It was therefore concluded that the reference point utilized by individuals in evaluating sources and messages does not contaminate the effect of open- or closed-mindedness upon this task.

It should be noted that, on the basis of the principle of cognitive congruity as outlined by Osgood and Tannenbaum (1955), the greater D^2 scores in the message reference condition might have been predicted.

The differential capacity of open and closed individuals to distinguish between and evaluate the substantive content and the source of a message, as substantiated by the results of this study, suggests some implications for the study of cognitive balance, source credibility, and reference group influences.

Open and closed individuals may differ in their tolerance of incongruity between their cognitions of the source and the message, and in the magnitude of the pressures which they experience to resolve such incongruity. Inasmuch as closed individuals are less able to differentiate between the source and message and to evaluate each on its own merits, they may be less tolerant of perceived incongruity between the two and hence experience greater pressures to reduce the incongruity than would open individuals.

Similarly, the influence of source credibility may operate differentially among open and closed individuals.

In the case of closed receivers, the credibility of the source may be a largely influential factor in determining whether or not the receiver's opinion will be changed as a result of a persuasive communication. Unless the message strongly contradicts the existing opinions of the receiver, a high credibility source may be more effective in changing the opinions of the receiver than a low credibility source.

In the case of open-minded receivers, capable of evaluating the source and the message independently, it may be that source credibility is not as influential in effecting opinion change. Opinion change in open individuals may be dependent largely upon the position advocated by the source and the receiver's evaluation of that position, independent of his evaluation of the source and his credibility. Source credibility may only take on importance when the open receiver is undecided or unable to evaluate the position advocated by the source. In this event, high credibility sources may be more effective than low credibility sources in effecting opinion change.

In the area of reference group influences, one of the more influential personality factors determining whether or not an individual will yield

to group pressures toward conformity may be that of the individual's degree of dogmatism. Closed individuals may be more responsive to group influences than are open individuals.

SUMMARY

The purpose of this study was to test the validity of the basic distinction between open and closed belief-disbelief systems suggested by Rokeach (1960) in his discussion of dogmatism.

It was hypothesized that the more open an individual's belief system, the greater should be his ability to distinguish between the message content and the message source and to judge each on its intrinsic merits.

Seventy-six subjects judged six source-statement pairs on a series of Semantic Differential scales. The sample was equally divided between open and closed subjects and these two groups were further equally divided between those subjects using the source as the reference point for judging the statements and those using the statements as the reference point for judging the source.

The results provided evidence that open and closed individuals do differ in their relative ability to differentiate between sources and messages and to evaluate them independently in a more or less realistic communication situation. Open individuals were found to be better able to do this than were closed individuals. This difference was found to exist regardless of the reference point used in evaluating the sources and messages.

REFERENCES

COLLIER, R. O., JR. Analysis of variance for correlated observations. *Psychometrika*, 1958, **23**, 223–236.

GREENHOUSE, S. W., & GEISSER, S. On methods in the analysis of profile data. *Psychometrika*, 1959, **24**, 95–112.

HOVLAND, C. I., & MANDELL, W. An experimental comparison of conclusion-drawing by the communicator and by the audience. *J. abnorm. soc. Psychol.*, 1952, **47**, 581–588.

HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Publ. opin. Quart.*, 1951, **15**, 635–650.

KELMAN, H. C., & HOVLAND, C. I. "Reinstatement" of the communicator in delayed measurement of opinion change. *J. abnorm. soc. Psychol.*, 1953, **48**, 327–333.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.

OSGOOD, C. E., & TANNENBAUM, P. H. The principle of congruity in the prediction of attitude change. *Psychol. Rev.*, 1955, **62**, 42–55.

ROKEACH, M. *The open and closed mind*. New York: Basic Books, 1960.

RYAN, T. A. Significance tests for multiple comparisons of proportions, variance, and other statistics. *Psychol. Bull.*, 1960, **57**, 318–328.

(Received May 6, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 64–67

MEMORIZING BY INDIVIDUALS AND GROUPS: A TEST OF THE POOLING-OF-ABILITY MODEL [1]

RONALD A. HOPPE [2]

Michigan State University

Perlmutter and DeMontmollin (1952) compared group learning of nonsense words (paralogs) with individual learning of the same nonsense words. Their procedure was to present a list of 19 two-syllable nonsense words and the task was to recall the words at the end of each presentation. Individuals and groups of three were presented with the same task. Learning curves were constructed for both the groups and the individuals. The groups started at a higher level of recall and remained at a higher level throughout the five learning trials. At the fifth trial the groups were averaging 18 of the 19 words, and the individuals were averaging 11 of the 19. The difference between groups and individuals was significant, and Perlmutter and DeMontmollin concluded that groups of three were superior to individuals in learning a list of nonsense syllables.

[1] The research was supported by the United States Office of Education, Grant 71-2141, Milton Rokeach, director. The author wishes to acknowledge the advice and criticism of Frank Restle who was a continuous help throughout the present research.

[2] Now at Miami University.

Perlmutter and DeMontmollin invoked an elaborate explanation involving interaction of the group members in order to account for the superiority of the group over the individual. However, the general results might be explained on the basis that individuals give only the words they separately remember, whereas group performance reflects the pooled learning of the members, who can be thought of as remembering separately. Perlmutter and DeMontmollin, at the time of their study, did not have a model available to test this hypothesis.

Lorge and Solomon (1955) offered a mathematical model of group problem solving to explain the results of Shaw (1932) who found that more three-member groups solved eureka-type problems than did individuals. According to the Lorge-Solomon model the probability that at least one of three individuals in a group will arrive at a solution is greater than the probability of success for one individual working separately. Let p_i be the probability that an individual will solve the problem within a given time. Then $(1 - p_i)$ is the probability that an individual will not arrive at the solution. If k individuals are working on the problem together, but each still has probability p_i of solving and the probabilities are independent, then the probability that none of them solves is $(1 - p_i)^k$. Therefore, the probability of the group arriving at the solution, p_g, is equal to $(1 - p_i)^k$. This model assumes no interaction which facilitates remembering in groups, but merely represents the pooling of ability of the members of the group.

The notion of pooling of abilities might also apply to Perlmutter and DeMontmollin's (1952) results. While the group remembered more nonsense words on each trial than did the individual, the group results may have been a pooling of words remembered by each of the members, and not attributable to an further facilitation from the group.

The purpose of this study is to replicate Perlmutter and DeMontmollin's (1952) study and to test the hypothesis that the superiority of the group can be explained by the Lorge and Solomon (1955) pooling model.

To apply the Lorge and Solomon (1955) model to group recall, the following consideration must be made. Treat the probability of recalling a previously seen nonsense word within a given time just as Lorge and Solomon treat the probability of solving a problem within a given time. Let p_i = the probability that an individual will recall a given word. Then the probability of a group, p_g, of size k of recalling the word would equal $1 - (1 - p_i)^k$.

TABLE 1

OBTAINED AND PREDICTED MEANS OF PERLMUTTER AND DEMONTMOLLIN (1952)

Trial	Individual means	Obtained group means	Predicted group means
1	2.3	5.5	6.1
2	5.4	10.2	12.0
3	7.7	14.8	15.0
4	9.7	17.3	16.8
5	11.4	18.2	17.8

To predict Perlmutter and DeMontmollin's (1952) results directly from this model it would be necessary to know the individual and group results of recalling each of the words in their list, but data were not available. However, a prediction of their results is possible if it is assumed that each of the words in the list has an equal probability of being remembered. Making this assumption the proportion of the entire list recalled by the group, p_{tg}, can be estimated from the proportion of the entire list recalled by individuals, p_{ti}, by the following formula: $p_{tg} = 1 - (1 - p_{ti})^k$. It follows that the predicted group mean, $\bar{G}$, is found by multiplying p_{tg} by the number of words on the list.

Using the formula mentioned above, Perlmutter and DeMontmollin's (1952) group results were predicted. Only a part of Perlmutter and DeMontmollin's results was used since they had the following balanced design: two groups of subjects were tested both as groups and individuals. Group A was given five recall trials as individuals on a list of 19 nonsense words, and then five recall trials as groups on a different but comparable list. Group B was given a list to recall as a group followed by another to recall as individuals. We used the individual results of Group A to predict the group results of Group B. Table 1 describes the predicted and obtained means for each trial. It can be seen that the obtained and predicted results are quite close.

The present experiment was concerned with testing the Lorge and Solomon (1955) model in a group and individual recall setting similar to that of Perlmutter and DeMontmollin (1952). The present experiment differed from Perlmutter and DeMontmollin's in that (*a*) the number of words to be recalled was reduced from 19 to 8; (*b*) both individual and groups were given nine instead of five trials; (*c*) more individuals and groups were tested than by Perlmutter and DeMontmollin.

METHOD

Subjects

The subjects were 65 voluntary introductory psychology students: 41 females and 24 males, divided into 14 groups of 3, where each group contained only members of the same sex; and 23 subjects tested individually. Each group was tested separately and the individuals were tested either singly or in pairs.

Procedure

Both the groups and individuals were told that they would be presented with eight nonsense words to recall. They were told to attempt to learn the words as quickly as possible. Both were told to record the recalled words on a pad of paper in front of them, one word per page, and after they had turned the page they were not allowed to turn back. Each individual recorded his own words, and in the groups the subject who sat in the middle was asked by the experimenter to record his own words and the other members' recalled words. The groups were also told to discuss each recalled word.

Occasionally a group asked if they could divide the task with each member taking a certain proportion of the words. To this the experimenter replied, "No," and that this would give results of three individuals instead of group results, and the experiment was concerned with group results. The experimenter observed no apparent division of the task.

After the instructions were read, both individuals and groups were presented the stimuli in prearranged random orders. The stimuli, constructed by combining three-letter nonsense syllables from Hilgard's (1951) list, were: BEMAW, HETIG, JALOB, KUQOM, LIPUF, MUWIY, SOHUJ, and ZADEC.

Groups and individuals were given nine trials with each of the above words presented once each trial. The words were presented in a different random order each trial and, furthermore, the order of the orders was randomized across individuals and groups. Each word was presented for 5 seconds followed by a 5-second interval in which no word was present. After the last word had been presented the subjects were given 2 minutes in which to recall the words. At the end of this period both the groups and individuals were told to mark an "X" on the next blank page, turn it, and await the next trial. At the end of the ninth trial the experimental session was concluded.

RESULTS AND DISCUSSION

The group results were predicted from the individual results, using the Lorge and Solomon (1955) model in the same manner as Perlmutter and Demontmollin's (1952) results were predicted. The assumption that each word had an equal chance of being recalled seemed tenable since the order of presentation of the words was randomized and the words were about equally meaningful. Figure 1 presents the mean number of words recalled correctly per trial compared with the predicted mean. An inspection of the figure suggests that the predicted and obtained results are quite close except for the first trial. The mean total number of words recalled by groups on all nine trials was 64.07 and the predicted mean total was 65.73. The difference of 1.66 was not significant ($t = 1.70$, $df = 13$, $p > .10$). The difference between the obtained and predicted results for each trial was also tested, using the t test. The only difference which was significant was the difference on the first trial ($t = 4.33$, $df = 13$, $p < .01$).[2]

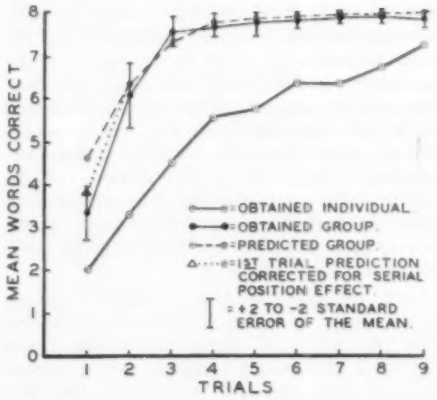

FIG. 1. Learning curves for individuals and groups, and group performance predicted from pooling of individual learning.

The difference on the first trial indicates an apparent weakness in the assumption that all words are equally likely to be remembered. While the positions of the words were randomized on each trial, a serial position effect could still occur on the first trial. If so, the assumption that each word has an equal chance of being recalled is

[2] There is no adequate statistic available to test the closeness of the predictions. The t test was used but in so doing the variance of the predictions was assumed to be zero and the predicted mean was treated as a parameter. The predictions, since they were based on the individual results, are variable, but there was no adequate estimate of this variance. The assumption of zero variance of the predictions was overly stringent for the purpose of this study. While it is possible that the difference on the first trial would not be significant if the variance of the predictions was known, the difference is still large enough to be concerned with, and suggests a flaw in the theory.

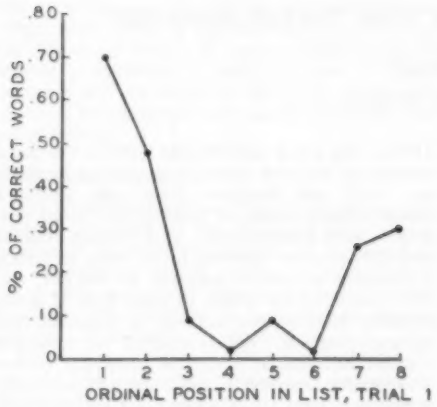

FIG. 2. First trial serial position effect obtained from individuals.

not tenable. Figure 2 shows the proportion of correct recalls by individuals at each of the eight serial positions on Trial 1. It can be seen that the first word presented was recalled by nearly 70% of the individuals while the words in the middle of the first presentation were recalled by less than 10% of the individuals, indicating a strong serial position effect.

What seemed to have happened on the first trial was that each of the group members tended to learn the *same* words. The pooling of abilities model with the assumption of each word having an equal chance of being recalled does not take this into account. The model pools the ability of all members with a disregard to which words were learned. An extreme example of this would be if all members of the groups and all individuals learned only the first and last words on the list. In this case the group would perform no better than the individual, but the prediction of the group's performance would be higher than the actual performance.

Because of the serial position effect, a new predicted mean, $\bar{G}$, for the first trial was made, based on the proportion of individuals recalling the words in each of the eight positions. The formula, $\hat{p}_g = 1 - (1 - \hat{p}_i)^k$, was applied separately for the words of each position. The new prediction of the mean number of words recalled, $\bar{G}$, was made by summing the predicted proportions of the words recalled by the group in each position, $\bar{G} = \Sigma\, \hat{p}_g$. The new prediction of 3.82 correct responses on Trial 1 was found not to differ significantly from the obtained value of 3.36 ($t = 1.49$, $df = 13$, $p > .10$).

From the above results it can be seen that the serial position effect can cause a discrepancy between the predicted and obtained group performance when the items in the list are lumped. The predictions of Perlmutter and DeMontmollin's (1952) results (Table 1) are somewhat high on early trials and low on later trials. If Perlmutter and DeMontmollin used the same order on each trial, the early high and later low predictions might be expected. They did not report whether or not they used the same order on each trial. In the present experiment with a different order of presentation of words on each trial, most of the serial position effect is likely to be cancelled out after the first trial.

The results of the present experiment suggest that one need not introduce a concept of group facilitation or interaction in order to explain the group learning of nonsense words.

SUMMARY

Sixty-five subjects were given eight nonsense words to learn. Forty-two of the subjects learned the task in groups of three, and the remaining 23 learned the task as individuals. The groups' results were predicted, using the Lorge and Solomon (1955) pooling model. The only prediction which differed significantly from the obtained was that of the first trial. When the serial position effect which occurred on the first trial was taken into account, the prediction no longer differed significantly from the obtained. The data of this study cannot be said to reject the Lorge and Solomon model and indicates that a concept of group facilitation need not be introduced to explain the groups' performance.

REFERENCES

HILGARD, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 517–567.

LORGE, I., & SOLOMON, H. Two models of group behavior in the solution of eureka-type problems. *Psychometrika*, 1955, **20**, 139–148.

PERLMUTTER, H. V., & DEMONTMOLLIN, G. Group learning of nonsense syllables. *J. abnorm. soc. Psychol.*, 1952, **47**, 762–769.

SHAW, MARJORIE E. Comparison of individuals and small groups in the rational solution of complex problems. *Amer. J. Psychol.*, 1932, **44**, 491–504.

(Received May 10, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 68–71

CONVERGENT VALIDATION OF RISK TAKING MEASURES [1]

PAUL SLOVIC [2]

University of Michigan

An increasing amount of recent research has attempted to relate various measures of what is thought to be a general psychological trait, called utility for risk or risk taking propensity, to other personality variables. Before the findings of these studies can be integrated into a comprehensive theory of risk taking behavior, the several different measures which they employ must demonstrate what Campbell and Fiske (1959) call convergent validity. This study attempts to provide evidence about such validity by determining whether the intercorrelations among several typical risk taking measures are significantly different from zero and sufficiently large to encourage further examination.

PROCEDURE

Eighty-two fraternity seniors were administered a battery of risk taking measures. The history and description of these measures are as follows:

Response Set Measures

Cronbach (1946) described three response sets which have recently been of particular interest to researchers seeking objective or stylistic response measures of risk taking. These are the tendency to guess on a doubtful test item rather than omit it (Cronbach termed this the gambling or caution versus incaution response set); the tendency to include many responses as members of a certain category when no limit is imposed (the inclusiveness set); the tendency to answer rapidly on timed tests, attempting to compensate for lower accuracy by completing more items (the speed versus accuracy set).

The gambling response set has been investigated as a personality mediated variable by Ziller (1957a, 1957b), who related it to occupational choice, and by Danielson (1957). Messick and Hills' (1960) "tab lifting" variable, an index of the amount of information a person requires before attempting to answer a test item, is also a measure of the gambling response set. The speed versus accuracy set has been studied by Johnson (1954), Winder and Wurtz (1954), and Block and Petersen (1955). The inclusiveness set has been related to personality by Gardner (1953) and Pettigrew (1958) who labeled it category width. Pettigrew asserted that broad categorizers have a tolerance for Type I errors. That is, they risk negative instances in an effort to include a maximum of positive instances. By contrast, narrow categorizers are willing to make Type II errors, excluding many positive instances by restricting their category ranges in order to minimize the number of negative instances. It should be noted that persons who sacrifice accuracy for speed and who are readily willing to gamble on an examination exhibit a willingness to risk Type I errors which is quite similar to that of the broad categorizer. Kogan and Wallach (1960) provided some evidence for the construct validity of these response set measures by relating a subject's willingness to risk errors to his Semantic Differential evaluation of risk laden concepts. Persons less willing to make errors rated such concepts as "risk," "stock market," "jet pilot," etc. as more hostile, cold, and tense than did people who were not as cautious.

The following response set measures were used in this study:

Dot Estimation Test (speed versus accuracy). The subject is given a short time in which to compare many pairs of dot filled squares. He is asked to indicate, for each pair, the square with the most dots in it. His score is the number of pairs attempted in the time allowed. A high score is hypothesized to be indicative of high risk taking.

Word Meanings (inclusiveness or category width). The subject is asked to check those of a group of 15 terms which belong in a stated category. The terms are quite ambiguous and it is predicted that high risk takers will check more terms than will low risk takers.

Test Risk (gambling set). This test was adapted from a method used by Coombs, Milholland, and Womer (1956) to determine the amount of knowledge an individual would have to have before he would guess on an examination in which guessing is penalized. The test consisted of 43, four alternative, multiple-choice, vocabulary items. The subject was told that he would receive one point for every wrong alternative which he correctly identified as wrong, but that he would lose three points for marking the right answer as being wrong. He was told that he could mark as many alternatives wrong for each item as he wished. After the test was completed, the subject was asked to look back at his answer sheet and attempt to guess the correct answer to each item out of alternatives which he had previously not marked. The number of those guesses which were correct was termed his conventional score. The extent to which

[1] Material included in this report stems in part from a project sponsored by the Personnel Laboratory, Wright Air Development Center, Air Research and Development Command, Lackland Air Force Base, Texas.

[2] The author wishes to thank Warren Norman and Clyde Coombs for their advice and suggestions in carrying out this study, and Cecil Mullins for permission to use his Dot Estimation, Word Meanings, and Self-Crediting tests.

this conventional score exceeds the subject's theoretical score (which is his expected number of correct guesses by chance) is assumed indicative of the amount of knowledge he really had about the questions but was not willing to risk using. Presumably, the more knowledge not used, the lower his risk taking propensity.

Questionnaire Measures

Life Experience Inventory. The Life Experience Inventory, developed by Torrance and Ziller (1957), consists of 38 items hypothesized to be related to the development of risk taking tendencies. These items cover sports and other recreational activities involving risk, reactions to competition, financial risks, risk in interpersonal relations, etc.

Job Preference Inventory. The Job Preference Inventory is an eight-item Guttman scale developed by Williams (1960) to measure risk taking in an industrial setting. People classified as high risk takers are those who state preferences for jobs in which they (*a*) are almost always on their own, (*b*) have to make many decisions by themselves, (*c*) could be either highly successful or a complete failure, etc.

Experimental Gambling Measures

Studies by Scodel, Minas, and Ratoosh (1959), Liverant and Scodel (1960), and Atkinson, Bastian, Earl, and Litwin (1960) have investigated the personality correlates of risk taking where risk taking was assessed by noting the subject's preferences and dislikes for gambles differing in probability of winning. Coombs and Pruitt (1960) argue that in addition to probability preferences, variance preferences are indicative of an individual's liking for risk. The variance of a bet reflects the amount of deviation of its possible outcomes from the average amount of money to be won or lost by playing the bet. For example, the bet $\frac{1}{2}$ to win \$5.00-$\frac{1}{2}$ to lose \$5.00 has a larger variance than the bet $\frac{1}{2}$ to win \$1.00-$\frac{1}{2}$ to lose \$1.00, although both have an expected value of zero dollars. It is evident that greater risk is concomitant with greater variance, when expectation and probability of winning and losing are held constant.

Bet Preference Test. Subjects rank ordered sets of four, two-outcome bets according to their preferences for playing them. No bets were actually played, however. These rankings were later decomposed into order relationships among pairs of bets. The bets had probabilities of winning ranging from 1/10 to 9/10 and five variance levels between .09 and 90.0. All bets had an expected value of zero dollars. A measure of variance preferences was obtained by noting the proportion of times the subject preferred the higher variance bet of a pair in which both gambles had the same probabilities of winning. A high proportion is hypothesized to be indicative of high risk taking. Similarly, a probability preference index was constructed by recording the proportion of times that the subject preferred the bet with lower probability of winning in a pair of bets which had the same variance. A high score on this index is indicative of a preference for long shots. Zero expected value bets with high probability of winning carry with them the possibility of a large loss. Long shot bets afford greater opportunities to lose but the amount of money that may be lost is relatively small. Therefore, there is no unequivocal hypothesis concerning the relationship between probability preferences among these bets and liking for high risk. It will be interesting, however, to compare those preferences with scores on the other risk taking measures.

Self-Crediting Test. The subject is shown a relatively easy vocabulary question and told that he will be given a test containing more items of similar difficulty. He is asked to set the number of points (between 1 and 10) which he wants every question on this test to be worth. The subject receives this amount of points for each question which he correctly answers. If he answers incorrectly, he loses this amount of points. After the first 10-item test he is told that the next test will be a more difficult one and he is to reset the number of points he wishes to give these harder questions. In all, four tests are given, and before each the subject is asked to reset the value of the items on the next, harder test. Holding knowledge constant, it is hypothesized that the high risk taker will want the next test to be worth more points than the low risk taker does. This test is really a variance preference measure. By adjusting the number of points at stake the subject is actually controlling the variance of the situation.

In addition to the above tasks, subjects were asked to rate their fellow fraternity brothers on a bipolar trait of general willingness to take risks. The A pole was defined as, "Loves to take risks. A daredevil." The B pole was labeled, "Cautious. Does not like to take chances. Avoids risky situations." Each rater was required to choose one-third of the rating group as being most like the A pole of the trait and one-third as being like the B pole with the remaining one-third designated as neither A nor B.

RESULTS AND DISCUSSION

Table 1 presents the intercorrelations among the risk taking measures. The variables (with the exception of the probability preference measure) have been reflected so that a high score is always considered indicative of high risk taking. The correlations with the Self-Crediting Test are partial correlations with the number of points risked on the test holding knowledge (the number of correct answers) constant.

Convergent validity requires that all correlations, except those with probability preferences, be significantly positive. The results show that only 5 correlations out of 28 reach significance in the predicted direction and none of these correlations exceeds .34. Another 10 of these correlations are negative, 2 significantly so. Even the response set tests which are quite similar in rationale and design fail to correlate with one another.

TABLE 1

INTERCORRELATIONS AMONG RISK TAKING MEASURES
($N = 82$)

Variable	1	2	3	4	5	6	7	8
Response sets								
1 Dot Estimation								
2 Word Meanings	−.17							
3 Test Risk	.16	.05						
Questionnaires								
4 Life Experience Inventory	.05	.27**	−.04					
5 Job Preference Inventory[a]	.07	−.14	−.19	−.06				
Gambling preferences								
6 Self-Crediting Test	−.08	.19*	−.24*	.05	.09			
7 Variance preferences	.32**	.03	−.07	.23*	.07	.04		
8 Probability preferences	.16	−.03	−.07	−.03	−.35*	−.20	−.17	
Ratings								
9 Risk rating	.05	.00	−.24*	.34**	.10	−.02	.02	.18*

Note.—Values of p (all tests, with the exception of those used with probability preferences) are one-tailed.
[a] $N = 55$.
* $p < .05$.
** $p < .01$.

With respect to probability preferences, there seems to be a slight tendency for people scoring in the high risk direction on the various tests to prefer bets with the higher probability of winning and lesser probabilities of a large loss.

It is quite obvious that these risk taking measures do not demonstrate convergent validity. Our findings are thus in agreement with those of Wallach and Kogan (1961) who, finding no relationship between willingness to risk errors on a judgmental task and scores on a probability preference measure, concluded that attempts to conceptualize various judgmental and cognitive processes in decision and risk terms must be considered premature.

The implications of the present study for the existence and measurement of a general risk taking trait are (*a*) none or only a few of the variables analyzed actually measure the trait; or (*b*) willingness to take risks may not be a general trait at all but rather one which varies from situation to situation within the same individual. In either case, future research must carefully consider the problem of adequately defining and assessing risk taking behavior.

SUMMARY

Eighty-two subjects were administered a battery of response set, questionnaire, experimental gambling, and peer rating measures of risk taking tendencies. The intercorrelations among these measures were generally not significant, indicating a lack of convergent validity.

REFERENCES

ATKINSON, J. W., BASTIAN, J. R., EARL, R. W., & LITWIN, G. H. The achievement motive, goal setting, and probability preferences. *J. abnorm. soc. Psychol.*, 1960, **60**, 27–36.

BLOCK, J., & PETERSEN, P. Some personality correlates of confidence, caution, and speed in a decision situation. *J. abnorm. soc. Psychol.*, 1955, **51**, 34–41.

CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.*, 1959, **56**, 81–105.

COOMBS, C. H., MILHOLLAND, J. E., & WOMER, F. B. The assessment of partial knowledge. *Educ. psychol. Measmt.*, 1956, **16**, 13–37.

COOMBS, C. H., & PRUITT, D. G. Components of risk in decision making: Probability and variance preferences. *J. exp. Psychol.*, 1960, **60**, 265–277.

CRONBACH, L. J. Response sets and test validity. *Educ. psychol. Measmt.*, 1946, **6**, 475–494.

DANIELSON, L. E. Gambling proneness: Its measurement and expression in examination situations. Unpublished doctoral dissertation, University of Michigan, 1957.

GARDNER, R. W. Cognitive styles in categorizing behavior. *J. Pers.*, 1953, **22**, 214–233.

JOHNSON, L. C. Speed and confidence of judgment as psychological variables. Stanford University, Department of Psychology, 1954. (Contract Nonr 225-01)

KOGAN, N., & WALLACH, M. Certainty of judgment and the evaluation of risk. *Psychol. Rep.*, 1960, **6**, 207–213.

LIVERANT, S., & SCODEL, A. Internal and external control as determinants of decision making under conditions of risk. *Psychol. Rep.*, 1960, **7**, 59–67.

MESSICK, S., & HILLS, J. Objective measurement of personality: Cautiousness and intolerance of am-

biguity. *Educ. psychol. Measmt.*, 1960, **20**, 685–698.

PETTIGREW, T. F. The measurement and correlates of category width as a cognitive variable. *J. Pers.*, 1958, **26**, 532–544.

SCODEL, A., MINAS, J., & RATOOSH, P. Some personality correlates of decision making under conditions of risk. *Behav. Sci.*, 1959, **4**, 19–28.

TORRANCE, E. P., & ZILLER, R. C. Risk and life experience: Development of a scale for measuring risk taking tendencies. *USAF Personnel Train. Res. Cent. res. Rep.*, 1957, No. 57-23.

WALLACH, M. A., & KOGAN, N. Aspects of judgment and decision making: Interrelationships and changes with age. *Behav. Sci.*, 1961, **6**, 23–36.

WILLIAMS, L. K. The measurement of risk taking propensity in an industrial setting. Unpublished doctoral dissertation, University of Michigan, 1960.

WINDER, C. L., & WURTZ, K. R. A study of personality correlates of judgment behavior. Stanford University, Department of Psychology, 1954. (Contract Nonr 225-01)

ZILLER, R. C. A measure of the gambling response set in objective tests. *Psychometrika*, 1957, **22**, 289–292. (a)

ZILLER, R. C. Vocational choice and utility for risk. *J. counsel. Psychol.*, 1957, **4**, 61–64. (b)

(Received May 27, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 1, 71–74

SOCIAL SCHEMAS AND THE RECONSTRUCTION OF SOCIAL OBJECT DISPLAYS FROM MEMORY

JAMES L. KUETHE

Johns Hopkins University

In a recent study it was demonstrated by Kuethe (1962) that there are basic social schemas that have high commonality for subjects. The investigation showed that when human figures, cut from felt, are placed on a felt board, the response is usually highly organized and similar to responses made by other individuals from the same population. The specific content of the figures in a set determined which schema was prepotent for individuals using that set.

Examples of schemas with high commonality were seen in the tendency of subjects to place a child nearer to a woman than to a man and to place a dog nearer to a man than to a woman. There was the basic social schema that "people belong together" which was manifested in the strong tendency of subjects to group human figures to a greater extent than nonhuman figures. Subjects who were allowed to place a set of figures containing a man, a woman, and two rectangles "wherever they wanted" grouped the human figures to a greater extent than the rectangles. The same result was obtained with three men figures and three rectangles. This schema provides the concept that human figures go together and nonhuman figures should not be allowed to intervene. For example, while the child was often placed between the man and woman figures, the dog was very rarely placed in that position.

This earlier investigation (using Bartlett's, 1932, method of successive reconstruction) showed that social schemas also function as social response sets in that they can introduce constant errors into judgments about the distance between social objects. When subjects viewed two rectangles a fixed distance apart they were quite accurate in replacing the rectangles in their original position on the field. However, with a set of man and woman figures there was the marked tendency to replace the figures too close together. The schema that man and woman belong together interfered with the judgment process.

One objective of the present investigation was to determine whether or not social schemas other than the man-woman schema interfere with judgments of figures having specific content. A specific hypothesis was that the strength of a schema, as determined by its frequency of use in a population, would directly relate to the magnitude of the error introduced by the schema into judgments involving objects that arouse the schema.

A separate objective, using a different technique, was to determine the manner in which a schema introduces error into the judgment process. This experiment was designed to determine whether a specific schema interferes with the original perception of a stimulus display or

whether interference occurs during the reconstruction of the stimulus display.

EXPERIMENT I

Method

The 50 subjects reconstructed the placement of the following pairs of stimulus figures: man and woman, two women, two men facing away from one another, two men facing each other, and two rectangles. Each pair of figures was presented on the display field with a 20-inch separation and the five pairs were presented in a different random order to each subject. (All measurements for both placement and recording of replacements were made from center of figure to center of figure.)

The field on which the figures were presented was a rectangle of blue felt, 2 yards long and 1.5 yards high, stretched on a wall of the experimental room. The figures, which were cut from yellow felt, could be placed in any part of the field. Because of the nap, one piece of felt will cling tightly to another. The advantage of this technique is that no mark is left on the field when the figures are removed. The large size of the field minimizes anchoring cues from the edges of the field.

All subjects were run individually and were male undergraduates at the Johns Hopkins University.

Results

The average separation of the reconstructed pairs was greatest for the two men figures facing away from each other (23.2 inches); the next greatest average separation upon replacement was for the rectangles (20.5 inches). It should be noticed that the average replacement of the two rectangles reveals that these objects are replaced with less error than any other pair of figures. The two woman figures were replaced closer together (18.3 inches) followed by the man and woman pair (17.1 inches). The two men facing one another were placed the closest together (16.0 inches). The concordance of all 50 subjects with this ranking was significant, $p < .05$. The five pairs of figures are shown in Figure 1. In Figure 1, the pairs of figures are ordered from left to right by decreasing mean magnitude of separation upon replacement.

EXPERIMENT II

Method

Each of the 100 subjects was given a set of felt figures consisting of a man, a woman, and a cat. The subject was told to place the figures on the felt field "in any manner he wanted." The relative position of the figures was recorded for each subject. The subjects performed the tasks in individual sessions.

Results

As may be seen in Figure 2, the order man-woman-cat was preferred and occurred 75 times. The order cat-man-woman occurred 10 times while the order man-cat-woman occurred 15 times. Also shown in Figure 2 are the frequencies with which 100 subjects used various orderings of a set containing man, woman, and dog figures. This last set has been discussed earlier by Kuethe (1962) and is presented here for comparison with the set containing a cat figure. The schema that places a dog near the man is not as strong as the schema that places a cat near a woman. However, each of these schemas is the schema with the highest commonality for its particular group of figures.

EXPERIMENT III

Method

The subjects in Experiment III had not participated in Experiments I and II. In individual sessions each of the 50 subjects viewed for 5 seconds the figures of the cat and the woman which had been placed on the felt field with a 20-inch separation by the experimenter. The subject then attempted to replace the figures on the field exactly were they had been. After the subject had placed the figures, the experimenter recorded their position.

The same procedure was used with the man and dog figures. Half of the subjects reconstructed the

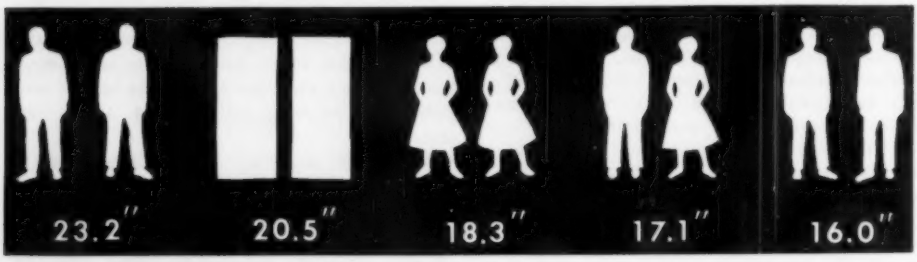

FIG. 1. Mean separation of the pairs of figures upon replacement after the figures were viewed with a 20-inch separation.

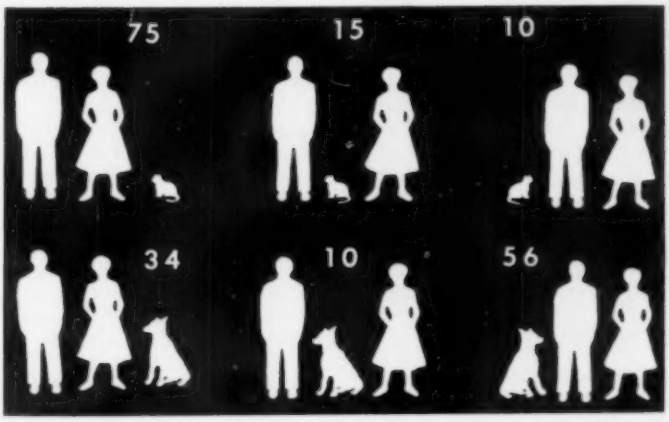

FIG. 2. Frequency with which 100 subjects used different orderings of the man, woman, and dog figures and the man, woman, and cat figures.

woman-cat display first and the rest of the subjects reconstructed the man-dog display first.

Results

The woman and cat figures were replaced closer together than the man and dog figures by 36 of the 50 subjects. This difference was significant by a sign test, $p < .01$. The woman and cat schema that has greater commonality among subjects than the schema that places a man and a dog together also introduces larger errors into judgments of the distance between these figures.

There is a rather subtle point that should be emphasized at this time; the original finding that the woman-cat schema has greater relative commonality than the man-dog schema does not necessarily guarantee the current result that there is a difference in the magnitude of error introduced by the two schemas. In the present experiment there were only two figures appearing at one time; there was no third figure to compete as in Experiment II.

EXPERIMENT IV

Method

In this experiment, subjects individually viewed for 5 seconds the man and woman figures which had been placed on the felt field with a 20-inch separation by the experimenter. The subjects then attempted to place *two rectangles* exactly where the human figures had been. The rectangles were the same size as the human figures. After the subject had placed the rectangles, the experimenter recorded their position.

The other part of this study used the same procedure in reverse. Each subject viewed for 5 seconds the two rectangles which had been placed on the field with a 20-inch separation. He then attempted to place the *male and female figures* exactly where the rectangles had been.

The order in which the two tasks were performed was counterbalanced; half of the subjects performed one task first and the rest performed the other task first.

Results

The man and woman figures were placed closer together than were the two rectangles by 44 of the 50 subjects. This difference was significant by a sign test, $p < .0001$.

This result becomes most meaningful when evaluated together with data from Experiment I. In that study it was found that the man and woman figures were replaced closer together than were the two rectangles. In the present experiment it was found that the man and woman figures were placed closer together than were the two rectangles even though the human figures were placed where the rectangles had been and the rectangles were placed where the human figures had been.

Subjects are relatively accurate when they use the rectangles to reproduce the relative position of either two rectangles or two human figures. Subjects err in placing man and woman figures too close together when they use these human figures to reproduce the relative positions of either two rectangles or two human figures. The schema that "people belong together" introduces errors in judgment during the reconstruction of a stimulus display rather than in the original perception of the stimulus display or in the

memory of the original display up until the time of reconstruction. If subjects originally perceived the man and woman figures as closer together than they really were, this would have been revealed when they placed the rectangles where they believed the human figures had been.

When a subject places one human figure, social schemas cause a second human figure to be placed nearer to it than would be the case if the figures lacked social content. Kuethe (1962) has shown that human figures are grouped to a greater extent than nonhuman figures in a free response situation.

DISCUSSION

The schema that people belong together introduces error into judgments that use human figures to reconstruct stimulus displays from memory. The extent of the distortion is related to the specific content of a set of figures. The tendency to place a man figure and a woman figure together is stronger than the tendency to place two woman figures together. That this may relate to a concept of relative interaction between people is seen in that two men figures facing one another are replaced considerably closer than originally viewed while two men figures facing away from one another are replaced further apart than originally viewed. There is very little error in the replacement of two rectangles as these objects fail to arouse any social schema.

The social schemas investigated were discovered by a free response technique which permitted subjects to place the figures wherever they wanted. The fact that the schemas discovered by this method may be shown to distort judgments in a predictable manner is critical. The distortion of judgment shows that the schemas are fundamental patterns of social organization for subjects. The distortion of judgment would not have occurred if the schemas discovered by the free response technique were merely the result of subjects attempting to respond the way they believed was "correct," or the way "most people" would respond.

In the study of social and other schemas, idiosyncratic organizations occur. An idiosyncratic organization can be the result of either the subject rejecting the schema that others tend to use or of the subject failing to possess the schema. Intentional rejection of the popular schema could reflect an attempt to be creative by rejecting the obvious. Failure to possess the schema could be the result of a cultural difference or could indicate a disturbance of social thinking.

The results of the present study suggest a technique for the evaluation of idiosyncratic schemas. Since high commonality social schemas introduce distortion into judgments concerning social objects, an individual who possesses the schema (even though he rejected it in free response) might be expected to show errors of judgment consistent with the popular schema. The judgment task emphasizes accuracy rather than creativity. An individual who really lacked the popular schema would not show distortions of judgment consistent with the popular schema but rather would be more likely to show distortions consistent with his idiosyncratic schema. The same situation may be seen when a subject does not use a "father and son" theme in response to the TAT card that usually elicits this schema. The interpretation would be quite different if it were known that the subject had decided the popular theme was too obvious and he was trying to be creative than if it could be shown that there was an actual disturbance or absence of the popular schema that an older man and a younger man are father and son.

SUMMARY

Social schemas discovered by a free response technique were found to distort reconstructions of the distance between pairs of figures. The specific content of the figures determined the magnitude of the distortion. In addition it was shown that the distortion of judgment occurs at the time of reconstruction rather than in the original perception or in retention.

Distortion is determined by the objects used in reconstruction, not by the objects originally viewed.

Discussion included the problem of analyzing whether the use of an idiosyncratic organization reflects a creative attempt to avoid the obvious or is indicative of a disturbance of the popular schema.

REFERENCES

BARTLETT, F. C. *Remembering: A study in experimental and social psychology.* New York: Macmillan, 1932.

KUETHE, J. L. Social schemas. *J. abnorm. soc. Psychol.*, 1962, **64**, 31–38.

(Early publication received November 10, 1961)

The Siege of Leningrad

LEON GOURE

Foreword by Merle Fainsod. From August 1941 to January 1944 the city of Leningrad was under siege, subject to constant German artillery fire and aerial bombardment. Almost one million of the city's inhabitants died of starvation, cold, and disease. Severe shortages of virtually all supplies, including water and fuel, contributed to the unrelenting misery of the people, and an almost total breakdown of city utilities and services was experienced. These conditions placed the political, administrative, and social systems of the city under extraordinary stress. Yet the Soviet authorities managed not only to maintain administrative order and control, but to mobilize the people into playing an active part in the city's defense. How this was accomplished is told here for the first time, providing the reader with a detailed account of one of the most dramatic and tragic episodes of the Second World War. Illustrated. $6.95

Order from your bookstore, please

STANFORD UNIVERSITY PRESS

PSYCHOLOGICAL TEST SPECIALISTS

Exclusive publishers and distributors of selected psychological tests

NEW!!

Quick Test (QT)

A very quick and satisfactorily reliable screening test of intellectual ability (6 to 10 minutes total time for all three forms). Words are defined non-verbally by choice among drawings. Suitable for any person (no matter how physically disabled) who can signal a yes-no response in any way. Covers intelligence levels from two years to superior adult. Excellent for building rapport, and for quick testing before and after any procedure.
$8.50 for set of materials

Memory-For-Designs Test (MFD)

Simple drawing test of perceptual-motor coordination, serving as extremely sensitive detector of brain injury of many types.
$8.50 for set of materials

NEW!!

Group Personality Projective Test (GPPT)

Multiple-choice questions about ambiguous stick-figures make possible rapid and effective measurement of significant, factor-analytically isolated aspects of personality.
$13.50 for set of materials

Id-Ego-Superego Test (IES)

Subtests give independent estimates of personality components from a modified psychoanalytic point of view.
$25.50 for set of materials

Kahn Test of Symbol Arrangement (KTSA)

Uses symbolic plastic objects which are arranged and sorted; validly discriminates among normals, neurotics, brain-injured, and psychotics, and indicates accompanying personality dynamics.
$25.00 for set of materials

Full-Range Picture Vocabulary Test (FRPV)

Brief test of intelligence in two forms using the same 16 plates. Words are defined by choosing pictures. Valuable for intelligence screening, and with the physically handicapped.
$7.50 for set of materials

Geist Picture Interest Inventory (GPII)

Provides for more valid expression of interests by choices among alternative pictured activities. Both qualitative and quantitative interpretations are possible.
$9.00 for set of materials

Proverbs Test (PT)

Group and individual forms determine similarity of thought processes to those of various normal and clinical groups. Highly sensitive indicator of psychotic processes.
$4.50 for kit

Vocational Apperception Test: Advanced Form (VAT:ADV)

Suggestive series of plates showing male and female college-related occupational activities, accompanied by explanation of use and interpretation as a projective test.
$7.25 for set of materials

Famous Sayings (FS)

Agreement with proverbs, aphorisms, and folk sayings is analyzed to determine personality structure. Particularly useful in personnel selection.
$4.50 for kit

It Scale For Children (ITSC)

An experimental picture scale for measuring sex role preference in children.
$15.00 for set of materials

Kahn Intelligence Tests: Experimental Form (KIT:EXP)

Relatively culture-free intelligence tests requiring no reading or writing. Supplementary scales for assessment of intelligence of blind, ability in concept formation, recall, and motor coordination. One scale can be administered by sign language.
$20.00 for set of materials

PSYCHOLOGICAL TEST SPECIALISTS

Box 1441 MISSOULA, MONTANA